

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

in - protein search, using sw model

November 25, 2002, 09:59:40 ; Search time 36 seconds
(without alignments)
736.580 Million cell updates/sec

US-09-785-738A-2

score: 1097

a: 1 MENEFDGKTEVLEAWYND.....KEFLAKLLKSGENQAVEGF 199

table: BLOSUM62

Gapop 10.0 , Gapext 0.5

d: 908470 seqs, 133250620 residues

number of hits satisfying chosen parameters: 908470

n DB seq length: 0

n DB seq length: 2000000000

rocessing: Minimum Match 0%

Listing First 45 summaries

se. A Geneseq_101002.*

```
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Alt to.	Score	Query Match	Length	ID	Description
1	1097	100.0	199	AAE08929	Rice submergence-i
2	955.5	87.1	198	AAE08930	Rice submergence-i
3	862	78.6	200	AAE08934	Cotton submergence
4	833.5	76.0	199	AAE08930	Arabidopsis thalia
5	833.5	76.0	199	AAE08930	Arabidopsis thalia
6	833.5	76.0	438	AAE08930	Arabidopsis thalia
7	833.5	76.0	495	AAE08930	Arabidopsis thalia
8	833.5	76.0	512	AAE08930	Arabidopsis thalia
9	838	75.5	192	AAE08930	Arabidopsis thalia
10	814	74.2	559	AAU72532	Arabidopsis cell c

11	812	74.0	196	22	AAE08931	Tomato submergence
12	800	72.9	198	21	AAE08931	Arabidopsis thalia
13	789.5	72.0	940	21	AAE08931	Arabidopsis thalia
14	782.5	71.3	182	21	AAE08931	Arabidopsis thalia
15	782.5	71.3	182	21	AAE08931	Arabidopsis thalia
16	781	71.2	199	21	AAE08931	Arabidopsis thalia
17	779	71.0	210	22	AAE08933	Soybean submergence
18	776	70.7	175	21	AAE08933	Arabidopsis thalia
19	776	70.7	175	23	AAE08933	Arabidopsis thalia
20	751	68.5	181	21	AAE08933	Arabidopsis thalia
21	737.5	67.2	923	21	AAE08933	Tomato submergence
22	691	63.0	189	22	AAE08932	Arabidopsis thalia
23	690	62.9	187	21	AAE08932	Arabidopsis thalia
24	690	62.9	199	21	AAE08932	Arabidopsis thalia
25	684	62.4	187	21	AAE08932	Arabidopsis thalia
26	684	62.4	198	21	AAE08932	Arabidopsis thalia
27	665	60.6	180	21	AAE08932	Arabidopsis thalia
28	659	60.1	180	21	AAE08932	Arabidopsis thalia
29	609	55.5	184	22	AAE08936	Mouse submergence-
30	586	53.4	191	22	AAE08936	Human protein seq
31	586	53.4	206	21	AAE08936	Human cancer assoc
32	586	53.4	206	23	AAE08936	Human ovarian anti
33	586	53.4	209	22	AAE08936	Human protein seq
34	581.5	53.0	181	22	AAE08936	zebrafish submerge
35	581	53.0	179	22	AAE08936	Human protein sequ
36	568	51.8	149	21	AAE08936	Arabidopsis thalia
37	568	51.8	866	21	AAE08936	Arabidopsis thalia
38	558	50.9	173	22	AAE08936	Human shear stress
39	542.5	49.5	125	21	AAE08936	Arabidopsis thalia
40	542.5	49.5	125	21	AAE08936	Arabidopsis thalia
41	542.5	49.5	125	21	AAE08936	Arabidopsis thalia
42	539.5	49.2	187	22	AAE08935	Human submergence-
43	535	48.8	118	21	AAE08935	Arabidopsis thalia
44	534.5	48.7	125	21	AAE08935	Arabidopsis thalia
45	529	48.2	124	21	AAE08935	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAE08929
ID AAE08929 standard; Protein; 199 AA.

XX AAE08929;

AC AC

DT 15-NOV-2001 (first entry)

XX Rice submergence-induced early response protein SH2A.
DE Rice; SH2A; submergence induced early response gene; anaerobic response;
KW water logging tolerance; hypoxic environment; horticultural;
KW agricultural; medical; fermentation; cell culture industry;
KW genetic construct; transgenic plant.

OS Oryza sativa.

XX Key Location/Qualifiers

FT Modified-site 21..24 /label= CK2

FT /note= "Casein kinase II phosphorylation site"

FT Modified-site 49..51 /label= PKC

FT /note= "Protein kinase C phosphorylation site"

FT Region 64..67 /note= "Nuclear localisation signal"

FT Modified-site 64..72 /label= TYR

FT /note= "Tyrosine kinase phosphorylation site"

FT Modified-site 73..76 /label= CK2

FT /note= "Casein kinase II phosphorylation site"

FT Modified-site 94..97

/label= CK2
/note= "Casein kinase II phosphorylation site"
102..105
Modified-site
/label= CK2
/note= "Casein kinase II phosphorylation site"
107..115
Modified-site
/label= TYR
/note= "Tyrosine kinase phosphorylation site"
145..152
Region
/label= Destruction_box
/note= "The presence of a destruction box motif
indicates that the protein undergoes rapid
turnover, which is a means of regulating its
activity"
145..153
Modified-site
/label= TYR
/note= "Tyrosine kinase phosphorylation site"
189..192
Modified-site
/label= CK2
/note= "Casein kinase II phosphorylation site"

WO200161020-A2.

23-AUG-2001.

16-FEB-2001; 2001WO-1B00745.

18-FEB-2000; 2000US-0183572.

(CROP-) CROPDISEIGN NV.

Sauter MM, Lorbiecke R;

WPI; 2001-529915/58.

N-PSDB; AADI5647.

New genetic constructs, comprising the SH2A or SH2A-like genes, useful for altering the growth and adaptation of plants under hypoxic conditions, e.g. improving water logging tolerance or regulating anaerobic response in plants -

Example 2; Page 91-92; 107pp; English.

The present invention relates to genetic constructs, which comprise a nucleotide sequence for an SH2A or SH2A-like gene, or an SH2A or SH2A-like protein gene promoter, where the SH2A or SH2A-like gene is operably linked to a promoter sequence that directs expression of the nucleotide sequence and the SH2A or SH2A-like protein gene promoter is operably linked to a heterologous coding sequence. The SH2A gene, initially identified in rice plants, is a submergence-induced early response gene which is involved in the adaptation of plants to hypoxic conditions (such as those encountered during submersion). In addition, the SH2A gene product has also been found to belong to a family of highly conserved proteins which occur ubiquitously in eukaryotes. The genetic constructs and methods are useful for altering the growth and adaptation of plants under hypoxic conditions. In particular, these are useful for improving water logging tolerance and regulating anaerobic response in plants. These are particularly useful in horticultural, agricultural, medical, fermentation and cell culture industries. The present sequence is rice SH2A protein.

sequence 199 AA;

Y Match 100.0%; Score 1097; DB 22; Length 199;
Local Similarity 100.0%; Pred. No. 5.1e-112;
hes 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MENEFOGKTEVTEAWYMDSEDDQRLPHHREPKFIPVDKLTGLVSWRLNPDNWNVC 60

1 MENEFOGKTEVTEAWYMDSEDDQRLPHHREPKFIPVDKLTGLVSWRLNPDNWNVC 60

61 ENLKRIREARGYSVDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLGSGYFDVRD 120

Db 61 ENLKRIREARGYSVDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLGSGYFDVRD 120
QY 121 QNDOWIRIALKKGMIVLPGMYHRTLTLDNYIKAMRLFVGDPVWTPYNRPHDLPAK 180
Db 121 QNDOWIRIALKKGMIVLPGMYHRTLTLDNYIKAMRLFVGDPVWTPYNRPHDLPAK 180
QY 181 EFLAKLLKSGEQAQVEGF 199
Db 181 EFLAKLLKSGEQAQVEGF 199

RESULT 2

AAE08930 standard; Protein; 198 AA.

XX AAE08930;

XX 15-NOV-2001 (first entry)

XX Rice submergence-induced early response protein SH2B.

XX Rice; SH2A; submergence induced early response gene; anaerobic response;
water logging tolerance; hypoxic environment; horticultural;
agricultural; medical; fermentation; cell culture industry;
genetic construct; transgenic plant; EST S2993; expressed sequence tag.

OS Oryza sativa.

PW WO200161020-A2.

XX 23-AUG-2001.

XX 16-FEB-2001; 2001WO-1B00745.

XX 18-FEB-2000; 2000US-0183572.

(CROP-) CROPDISEIGN NV.

XX Sauter MM, Lorbiecke R;

DR WPI; 2001-529915/58.

XX N-PSDB; AADI5648.

XX New genetic constructs, comprising the SH2A or SH2A-like genes, useful for altering the growth and adaptation of plants under hypoxic conditions, e.g. improving water logging tolerance or regulating anaerobic response in plants -

Example 3; Page 93-94; 107pp; English.

The present invention relates to genetic constructs, which comprise a nucleotide sequence for an SH2A or SH2A-like gene, or an SH2A or SH2A-like protein gene promoter, where the SH2A or SH2A-like gene is operably linked to a promoter sequence that directs expression of the nucleotide sequence and the SH2A or SH2A-like protein gene promoter is operably linked to a heterologous coding sequence. The SH2A gene, initially identified in rice plants, is a submergence-induced early response gene which is involved in the adaptation of plants to hypoxic conditions (such as those encountered during submersion). In addition, the SH2A gene product has also been found to belong to a family of highly conserved proteins which occur ubiquitously in eukaryotes. The genetic constructs and methods are useful for altering the growth and adaptation of plants under hypoxic conditions. In particular, these are useful for improving water logging tolerance and regulating anaerobic response in plants. These are particularly useful in horticultural, agricultural, medical, fermentation and cell culture industries. The present sequence is a rice SH2A-like protein, SH2B, encoded by EST (expressed sequence tag) S2993.

Sequence 198 AA;

Query Match 87.1%; Score 955.5; DB 22; Length 198;
Best Local Similarity 85.3%; Pred. No. 1.7e-96;
Matches 168; Conservative 19. Mismatch

14-MAY-1999; 99US-0134370.
18-MAY-1999; 99US-0134769.
19-MAY-1999; 99US-0134941.
20-MAY-1999; 99US-0135124.
21-MAY-1999; 99US-0135353.
24-MAY-1999; 99US-0135629.
25-MAY-1999; 99US-0136021.
27-MAY-1999; 99US-0136392.
28-MAY-1999; 99US-0136782.
01-JUN-1999; 99US-0137222.
03-JUN-1999; 99US-0137528.
04-JUN-1999; 99US-0137502.
07-JUN-1999; 99US-0137724.
08-JUN-1999; 99US-0138094.
10-JUN-1999; 99US-0138540.
10-JUN-1999; 99US-0138847.
14-JUN-1999; 99US-0139119.
16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
17-JUN-1999; 99US-0139492.
18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140353.
23-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140695.
28-JUN-1999; 99US-0140823.
29-JUN-1999; 99US-0140991.
30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142920.
12-JUL-1999; 99US-0142977.
13-JUL-1999; 99US-0143542.
14-JUL-1999; 99US-0143624.
15-JUL-1999; 99US-0144005.
16-JUL-1999; 99US-0144085.
16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-0144325.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
19-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-0144335.
20-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144884.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145086.
21-JUL-1999; 99US-0145088.
22-JUL-1999; 99US-0145085.
22-JUL-1999; 99US-0145087.
22-JUL-1999; 99US-0145089.
23-JUL-1999; 99US-0145192.
23-JUL-1999; 99US-0145145.
23-JUL-1999; 99US-0145218.
23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145913.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
28-JUL-1999; 99US-0145951.
28-AUG-1999; 99US-0146386.
02-AUG-1999; 99US-0146388.
02-AUG-1999; 99US-0146389.
02-AUG-1999; 99US-0147038.
03-AUG-1999; 99US-0147204.
04-AUG-1999; 99US-0147302.
05-AUG-1999; 99US-0147192.
05-AUG-1999; 99US-0147260.
06-AUG-1999; 99US-0147303.
06-AUG-1999; 99US-0147416.
09-AUG-1999; 99US-0147493.
09-AUG-1999; 99US-0147935.
10-AUG-1999; 99US-0148171.
11-AUG-1999; 99US-0148319.
12-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-0148565.
13-AUG-1999; 99US-0148684.
16-AUG-1999; 99US-0149368.
17-AUG-1999; 99US-0149175.
18-AUG-1999; 99US-0149426.
20-AUG-1999; 99US-0149722.
20-AUG-1999; 99US-0149723.
20-AUG-1999; 99US-0149929.
23-AUG-1999; 99US-0149930.
23-AUG-1999; 99US-0149930.
25-AUG-1999; 99US-0150566.
26-AUG-1999; 99US-0150884.
27-AUG-1999; 99US-0151065.
27-AUG-1999; 99US-0151066.
27-AUG-1999; 99US-0151080.
30-AUG-1999; 99US-0151303.
31-AUG-1999; 99US-0151438.
01-SEP-1999; 99US-0151930.
01-SEP-1999; 99US-0152363.
07-SEP-1999; 99US-0153070.
10-SEP-1999; 99US-0153758.
13-SEP-1999; 99US-0154018.
15-SEP-1999; 99US-0154039.
16-SEP-1999; 99US-0154779.
20-SEP-1999; 99US-0155139.
22-SEP-1999; 99US-0155486.
23-SEP-1999; 99US-0155659.
24-SEP-1999; 99US-0156458.
28-SEP-1999; 99US-0156458.
29-SEP-1999; 99US-0156596.
04-OCT-1999; 99US-0157117.
05-OCT-1999; 99US-0157753.
06-OCT-1999; 99US-0157865.
07-OCT-1999; 99US-0158029.
08-OCT-1999; 99US-0158232.
12-OCT-1999; 99US-0158369.
13-OCT-1999; 99US-0159293.
13-OCT-1999; 99US-0159294.
13-OCT-1999; 99US-0159295.
14-OCT-1999; 99US-0159329.
14-OCT-1999; 99US-0159330.
14-OCT-1999; 99US-0159331.
14-OCT-1999; 99US-0159637.
14-OCT-1999; 99US-0159638.
18-OCT-1999; 99US-0159584.
21-OCT-1999; 99US-0160741.
21-OCT-1999; 99US-0160767.
21-OCT-1999; 99US-0160768.
21-OCT-1999; 99US-0160770.
21-OCT-1999; 99US-0160814.
21-OCT-1999; 99US-0160815.
22-OCT-1999; 99US-0160980.
22-OCT-1999; 99US-0160981.
22-OCT-1999; 99US-0160989.
25-OCT-1999; 99US-0161404.
25-OCT-1999; 99US-0161405.

99US-01451145.
23-JUL-1999;
99US-0145218.
23-JUL-1999;
99US-0145224.
26-JUL-1999;
99US-0145276.
27-JUL-1999;
99US-0145913.
27-JUL-1999;
99US-0145918.
27-JUL-1999;
99US-0145919.
28-JUL-1999;
99US-0145951.
02-AUG-1999;
99US-0146386.
02-AUG-1999;
99US-0146388.
02-AUG-1999;
99US-0146389.
03-AUG-1999;
99US-0147038.
04-AUG-1999;
99US-0147204.
04-AUG-1999;
99US-0147302.
05-AUG-1999;
99US-0147192.
05-AUG-1999;
99US-0147260.
06-AUG-1999;
99US-0147303.
06-AUG-1999;
99US-0147416.
09-AUG-1999;
99US-0147493.
09-AUG-1999;
99US-0147935.
10-AUG-1999;
99US-0148171.
11-AUG-1999;
99US-0148319.
12-AUG-1999;
99US-0148341.
13-AUG-1999;
99US-0148565.
13-AUG-1999;
99US-0148684.
16-AUG-1999;
99US-0149368.
17-AUG-1999;
99US-0149175.
18-AUG-1999;
99US-0149426.
20-AUG-1999;
99US-0149722.
20-AUG-1999;
99US-0149723.
20-AUG-1999;
99US-0149929.
23-AUG-1999;
99US-0149902.
23-AUG-1999;
99US-0149930.
25-AUG-1999;
99US-0150566.
26-AUG-1999;
99US-0150884.
27-AUG-1999;
99US-0151065.
27-AUG-1999;
99US-0151066.
27-AUG-1999;
99US-0151080.
30-AUG-1999;
99US-0151303.
31-AUG-1999;
99US-0151438.
01-SEP-1999;
99US-0151930.
07-SEP-1999;
99US-0152363.
10-SEP-1999;
99US-0153070.
13-SEP-1999;
99US-0153758.
15-SEP-1999;
99US-0154018.
16-SEP-1999;
99US-0154039.
20-SEP-1999;
99US-0154779.
22-SEP-1999;
99US-0155139.
23-SEP-1999;
99US-0155486.
24-SEP-1999;
99US-0155659.
28-SEP-1999;
99US-0156458.
29-SEP-1999;
99US-0156596.
04-OCT-1999;
99US-0157117.
05-OCT-1999;
99US-0157753.
06-OCT-1999;
99US-0157865.
07-OCT-1999;
99US-0158029.
08-OCT-1999;
99US-0158232.
12-OCT-1999;
99US-0158369.
13-OCT-1999;
99US-0159293.
13-OCT-1999;
99US-0159294.
14-OCT-1999;
99US-0159295.
14-OCT-1999;
99US-0159329.
14-OCT-1999;
99US-0159330.
14-OCT-1999;
99US-0159331.
14-OCT-1999;
99US-0159637.
18-OCT-1999;
99US-0159638.
18-OCT-1999;
99US-0159584.
21-OCT-1999;
99US-0160741.
21-OCT-1999;
99US-0160767.
21-OCT-1999;
99US-0160768.
21-OCT-1999;
99US-0160770.
21-OCT-1999;
99US-0160814.
21-OCT-1999;
99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
Query Match 76.0%; Score 833.5; DB 21; Length 199;
Best Local Similarity 76.5%; Pred. No. 4.3e-83;
Matches 143; Conservative 29; Mismatches 14; Indels 1; Gaps 1;
QY 6 QDGKTEVIAWYDDSEEDORLPHHREPKEFIPVDKLTGLVSWRLNPWNWENCENLKR 65
Db 6 KDGREEVIAWYDDSEEDORLPHHDKPKFELSLDKLAELGVLWSRLDADNYETDEDLKK 65
QY 66 IREARGYSYVDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLEGSGYFVDRDNDOW 125
Db 66 IREARGYSYMDFCVCEPKLPNYEVKVSFEEHLHTDEIRYCVAGSGYFVDRNEAW 125
QY 126 IRIALKKGGMIVLPAGMYHRTDLDNYIKAMRLFVGDPVWTPYNNRPHDLPARKEFLAK 185
Db 126 IRVWVKGGMIVLPAGIYHRTVDSDNYIKAMRLFVGPVWTPYNNRPHDLPARKEYIDN 185
QY 186 LLK-SEG 191
Db 186 FVKVNEG 192
RESULT 6
AAG45259
ID AAG45259 standard; Protein: 438 AA.
XX AAG45259;
AC AAG45259;
XX AAG45259;
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 56796.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 56796.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EPI033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.

30-APR-1999; 99US-0132407.
04-MAY-1999; 99US-0132484.
05-MAY-1999; 99US-0132485.
06-MAY-1999; 99US-0132486.
06-MAY-1999; 99US-0132487.
07-MAY-1999; 99US-0132863.
11-MAY-1999; 99US-0134256.
14-MAY-1999; 99US-0134218.
14-MAY-1999; 99US-0134219.
14-MAY-1999; 99US-0134221.
14-MAY-1999; 99US-0134370.
18-MAY-1999; 99US-0134768.
19-MAY-1999; 99US-0134941.
20-MAY-1999; 99US-0135124.
21-MAY-1999; 99US-0135353.
24-MAY-1999; 99US-0135625.
25-MAY-1999; 99US-0136021.
27-MAY-1999; 99US-0136392.
28-MAY-1999; 99US-0136782.
01-JUN-1999; 99US-0137222.
03-JUN-1999; 99US-0137528.
04-JUN-1999; 99US-0137502.
07-JUN-1999; 99US-0137724.
08-JUN-1999; 99US-0138094.
10-JUN-1999; 99US-0138540.
10-JUN-1999; 99US-0138847.
14-JUN-1999; 99US-0139119.
16-JUN-1999; 99US-0139452.
16-JUN-1999; 99US-0139453.
17-JUN-1999; 99US-0139492.
18-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140353.
23-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140695.
28-JUN-1999; 99US-0140823.
29-JUN-1999; 99US-0140991.
30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142920.
12-JUL-1999; 99US-0142977.
13-JUL-1999; 99US-0143542.
14-JUL-1999; 99US-0143624.
15-JUL-1999; 99US-0144005.
16-JUL-1999; 99US-0144085.
16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-0144325.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
19-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-0144335.
20-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144884.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145086.
21-JUL-1999; 99US-0145087.
21-JUL-1999; 99US-0145088.
21-JUL-1999; 99US-0145089.
22-JUL-1999; 99US-0145089.
22-JUL-1999; 99US-0145192.
23-JUL-1999; 99US-0145145.
23-JUL-1999; 99US-0145218.
23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145913.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
28-JUL-1999; 99US-0145951.
02-AUG-1999; 99US-0146386.
02-AUG-1999; 99US-0146388.
02-AUG-1999; 99US-0146389.
03-AUG-1999; 99US-0147038.
04-AUG-1999; 99US-0147204.
04-AUG-1999; 99US-0147302.
05-AUG-1999; 99US-0147192.
05-AUG-1999; 99US-0147260.
06-AUG-1999; 99US-0147303.
06-AUG-1999; 99US-0147416.
09-AUG-1999; 99US-0147493.
09-AUG-1999; 99US-0147935.
10-AUG-1999; 99US-0148171.
11-AUG-1999; 99US-0148319.
12-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-0148565.
13-AUG-1999; 99US-0148684.
16-AUG-1999; 99US-0149368.
17-AUG-1999; 99US-0149175.
18-AUG-1999; 99US-0149426.
20-AUG-1999; 99US-0149722.
20-AUG-1999; 99US-0149723.
20-AUG-1999; 99US-0149929.
23-AUG-1999; 99US-0149902.
23-AUG-1999; 99US-0149930.
25-AUG-1999; 99US-0150566.
26-AUG-1999; 99US-0150884.
27-AUG-1999; 99US-0151065.
27-AUG-1999; 99US-0151066.
27-AUG-1999; 99US-0151080.
30-AUG-1999; 99US-0151303.
31-AUG-1999; 99US-0151438.
01-SEP-1999; 99US-0151930.
07-SEP-1999; 99US-0152363.
10-SEP-1999; 99US-0153070.
13-SEP-1999; 99US-0153758.
15-SEP-1999; 99US-0154018.
16-SEP-1999; 99US-0154039.
20-SEP-1999; 99US-0154779.
22-SEP-1999; 99US-0155139.
23-SEP-1999; 99US-0155486.
24-SEP-1999; 99US-0155659.
28-SEP-1999; 99US-0156458.
29-SEP-1999; 99US-0156596.
04-OCT-1999; 99US-0157117.
05-OCT-1999; 99US-0157753.
06-OCT-1999; 99US-0157865.
07-OCT-1999; 99US-0158029.
08-OCT-1999; 99US-0158232.
12-OCT-1999; 99US-0158369.
13-OCT-1999; 99US-0159283.
13-OCT-1999; 99US-0159294.
13-OCT-1999; 99US-0159295.
14-OCT-1999; 99US-0159329.
14-OCT-1999; 99US-0159330.
14-OCT-1999; 99US-0159331.
14-OCT-1999; 99US-0159637.
14-OCT-1999; 99US-0159638.
18-OCT-1999; 99US-0159584.
21-OCT-1999; 99US-0160741.

?	21-OCT-1999;	99US-0160767.	PR	21-APR-1999;	99US-0130449.
?	21-OCT-1999;	99US-0160768.	PR	23-APR-1999;	99US-0130510.
?	21-OCT-1999;	99US-0160770.	PR	23-APR-1999;	99US-0130891.
?	21-OCT-1999;	99US-0160814.	PR	28-APR-1999;	99US-0131449.
?	21-OCT-1999;	99US-0160815.	PR	30-APR-1999;	99US-0132048.
?	22-OCT-1999;	99US-0160980.	PR	30-APR-1999;	99US-0132407.
?	22-OCT-1999;	99US-0160981.	PR	04-MAY-1999;	99US-0132484.
?	22-OCT-1999;	99US-0160989.	PR	05-MAY-1999;	99US-0132485.
?	25-OCT-1999;	99US-0161404.	PR	06-MAY-1999;	99US-0132486.
?	25-OCT-1999;	99US-0161405.	PR	06-MAY-1999;	99US-0132487.
?	25-OCT-1999;	99US-0161406.	PR	07-MAY-1999;	99US-0132863.
?	26-OCT-1999;	99US-0161359.	PR	11-MAY-1999;	99US-0134256.
?	26-OCT-1999;	99US-0161360.	PR	14-MAY-1999;	99US-0134218.
?	26-OCT-1999;	99US-0161361.	PR	14-MAY-1999;	99US-0134219.
?	28-OCT-1999;	99US-0161920.	PR	14-MAY-1999;	99US-0134221.
?	28-OCT-1999;	99US-0161992.	PR	14-MAY-1999;	99US-0134370.
?	28-OCT-1999;	99US-0161993.	PR	18-MAY-1999;	99US-0134768.
?	29-OCT-1999;	99US-0162142.	PR	19-MAY-1999;	99US-0134941.
?	29-OCT-1999;	99US-0162142.	PR	20-MAY-1999;	99US-0135124.
Query Match 76.0%; Score 833.5; DB 21; Length 438;					99US-0135353.
Best Local Similarity 76.5%; Pred No. 1.3e-82;					99US-0135629.
Matches 143; Conservative 29; Mismatches 14; Indels 1; Gaps 1;					99US-0136021.
6 QDGKTEVIAWYMDSEEDORLPHHREPKETIPVDKLTGELGYISWRLPDNNWENCLNR 65					99US-0136392.
245 KDGREVIQAWYMDSEEDQRLPHHKDPKEFLSLDKLAELGVLSWRLDADNYETDDELLK 304					99US-0136782.
66 IREARGSYVDICDVCPEKLPNYEKIKSFFPEHLHTDEIEYICLEGSGYFVDRDNDQW 125					99US-0137222.
305 IRESRGYSYMDCEVCPEKLPNYEKVKSFPEHLHTDEIEYICVAGSGYFVDRDNEAW 364					99US-0137528.
126 IRIALKKGMIVLPAGIYHRTLDITDNYIKAMRLFVGDVWTPYNNRPHDHPARKEFLAK 185					99US-0137502.
365 IRVWKKGMIVLPAGIYHRTVDSNYYIKAMRLFVGEFVWTPYNNRPHDHPARKEYIDN 424					99US-0137724.
186 LK-SEG 191					99US-0138094.
425 FVKVNEG 431					99US-0138540.
SULT 7					99US-0138847.
345258					99US-0139119.
AAG45258 standard; Protein; 495 AA.					99US-0139452.
AAG45258;					99US-0139453.
18-OCT-2000 (first entry)					99US-0139454.
Arabidopsis thaliana protein fragment SEQ ID NO: 56795.					99US-0139454.
Protein identification; signal transduction pathway; metabolic pathway;					99US-0139455.
hybridisation assay; genetic mapping; gene expression control; promoter;					99US-0139455.
termination sequence.					99US-0139456.
Arabidopsis thaliana.					99US-0139457.
EP1033405-A2.					99US-0139458.
06-SEP-2000.					99US-0139459.
25-FEB-2000; 2000EP-0301439.					99US-0139460.
25-FEB-1999; 99US-0121825.					99US-0139461.
05-MAR-1999; 99US-0123180.					99US-0139462.
09-MAR-1999; 99US-0123548.					99US-0139463.
23-MAR-1999; 99US-0125788.					99US-0139463.
25-MAR-1999; 99US-0126264.					99US-0139463.
29-MAR-1999; 99US-0126785.					99US-0139463.
01-APR-1999; 99US-0127462.					99US-0139463.
06-APR-1999; 99US-0128234.					99US-0139463.
08-APR-1999; 99US-0128714.					99US-0139463.
16-APR-1999; 99US-0129845.					99US-0139463.
19-APR-1999; 99US-0130077.					99US-0139463.

20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.	QY	6	ODGKTEVIEAWYMDSEEDQRLPHHREPKEFIPVDKLTGLGVISWRLNPDWNCENLKR	65
20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.	DB	302	KDREEVIQAWYMDSEEDQRLPHHDPKEFLSLDKLAELGVLWSRUDADNYETDEDLK	361
20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.	QY	66	IREARGSYVDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLEGSGYFVDRDQNDQW	125
21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.	DB	362	IRESRGYSYMDFCVCEPKLPNYEVKVSFEEHLHTDEIRYCVAGSGYFVDRDRNEAW	421
21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.	QY	126	IRIALKKGGMIVLPAGMYHRFTLTDNYIKAMRLFVGDVWTPYNNRPHDHLPARKEFLAK	185
21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.	DB	422	IRVWVKKGGMIVLPAGIYHRFTVDSNDNYIKAMRLFVGPVWTPYNNRPHDHLPARKEYIDN	481
22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.	QY	186	LLK-SEG 191	
22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.	DB	482	FVKVNEG 488	
22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.	RESULT 8			
22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.	AAG45257			
23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.	ID	AAG45257 standard; Protein; 512 AA.		
23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.	XX	AAG45257;		
23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.	XX	18-OCT-2000 (first entry)		
26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0161404.	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 56794.		
27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.	XX	Protein identification; signal transduction pathway; metabolic pathway;		
27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.	KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.	KW	termination sequence.		
28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.	XX	Arabidopsis thaliana.		
02-AUG-1999;	99US-0146380.	PR	26-OCT-1999;	99US-0161361.	OS			
02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161920.	XX	EP1033405-A2.		
03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161992.	FN	06-SEP-2000.		
04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161993.	XX	25-FEB-2000; 2000EP-0301439.		
04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.	PF	25-FEB-1999; 99US-0121825.		
05-AUG-1999;	99US-0147192.	PR			XX	05-MAR-1999; 99US-0123180.		
05-AUG-1999;	99US-0147260.	PR			PR	09-MAR-1999; 99US-0123548.		
06-AUG-1999;	99US-0147303.	PR			PR	23-MAR-1999; 99US-0125788.		
06-AUG-1999;	99US-0147416.	PR			PR	25-MAR-1999; 99US-0126264.		
09-AUG-1999;	99US-0147493.	PR			PR	29-MAR-1999; 99US-0126785.		
09-AUG-1999;	99US-0147935.	PR			PR			
10-AUG-1999;	99US-0148171.	PR			PR			
11-AUG-1999;	99US-0148319.	PR			PR			
12-AUG-1999;	99US-0148341.	PR			PR			
13-AUG-1999;	99US-0148565.	PR			PR			
13-AUG-1999;	99US-0148684.	PR			PR			
16-AUG-1999;	99US-0149368.	PR			PR			
17-AUG-1999;	99US-0149175.	PR			PR			
18-AUG-1999;	99US-0149426.	PR			PR			
20-AUG-1999;	99US-0149722.	PR			PR			
20-AUG-1999;	99US-0149723.	PR			PR			
20-AUG-1999;	99US-0149929.	PR			PR			
23-AUG-1999;	99US-0149902.	PR			PR			
23-AUG-1999;	99US-0149930.	PR			PR			
25-AUG-1999;	99US-0150566.	PR			PR			
26-AUG-1999;	99US-0150884.	PR			PR			
27-AUG-1999;	99US-0151065.	PR			PR			
27-AUG-1999;	99US-0151066.	PR			PR			

? 01-APR-1999; 99US-0127462.
? 06-APR-1999; 99US-0128234.
? 08-APR-1999; 99US-0128714.
? 16-APR-1999; 99US-0129845.
? 19-APR-1999; 99US-0130077.
? 21-APR-1999; 99US-0130449.
? 23-APR-1999; 99US-0130510.
? 23-APR-1999; 99US-0130891.
? 28-APR-1999; 99US-0131449.
? 30-APR-1999; 99US-0132048.
? 30-APR-1999; 99US-0132407.
? 04-MAY-1999; 99US-0132484.
? 05-MAY-1999; 99US-0132485.
? 06-MAY-1999; 99US-0132486.
? 06-MAY-1999; 99US-0132487.
? 07-MAY-1999; 99US-0132863.
? 11-MAY-1999; 99US-0134256.
? 14-MAY-1999; 99US-0134218.
? 14-MAY-1999; 99US-0134219.
? 14-MAY-1999; 99US-0134221.
? 14-MAY-1999; 99US-0134370.
? 18-MAY-1999; 99US-0134768.
? 19-MAY-1999; 99US-0134941.
? 20-MAY-1999; 99US-0135124.
? 21-MAY-1999; 99US-0135353.
? 24-MAY-1999; 99US-0135629.
? 25-MAY-1999; 99US-0136021.
? 27-MAY-1999; 99US-0136392.
? 28-MAY-1999; 99US-0136782.
? 01-JUN-1999; 99US-0137222.
? 03-JUN-1999; 99US-0137528.
? 04-JUN-1999; 99US-0137502.
? 07-JUN-1999; 99US-0137724.
? 08-JUN-1999; 99US-0138094.
? 10-JUN-1999; 99US-0138540.
? 10-JUN-1999; 99US-0138847.
? 14-JUN-1999; 99US-0139119.
? 16-JUN-1999; 99US-0139452.
? 16-JUN-1999; 99US-0139453.
? 17-JUN-1999; 99US-0139492.
? 18-JUN-1999; 99US-0139454.
? 18-JUN-1999; 99US-0139455.
? 18-JUN-1999; 99US-0139456.
? 18-JUN-1999; 99US-0139457.
? 18-JUN-1999; 99US-0139458.
? 18-JUN-1999; 99US-0139459.
? 18-JUN-1999; 99US-0139460.
? 18-JUN-1999; 99US-0139461.
? 18-JUN-1999; 99US-0139462.
? 18-JUN-1999; 99US-0139463.
? 18-JUN-1999; 99US-0139750.
? 18-JUN-1999; 99US-0139763.
? 21-JUN-1999; 99US-0139817.
? 22-JUN-1999; 99US-0139899.
? 23-JUN-1999; 99US-0140353.
? 23-JUN-1999; 99US-0140354.
? 24-JUN-1999; 99US-0140695.
? 24-JUN-1999; 99US-0140823.
? 28-JUN-1999; 99US-0140823.
? 29-JUN-1999; 99US-0140991.
? 30-JUN-1999; 99US-0141287.
? 01-JUL-1999; 99US-0141842.
? 01-JUL-1999; 99US-0142154.
? 02-JUL-1999; 99US-0142055.
? 06-JUL-1999; 99US-0142390.
? 08-JUL-1999; 99US-0142803.
? 09-JUL-1999; 99US-0142920.
? 12-JUL-1999; 99US-0142977.
? 13-JUL-1999; 99US-0143542.
? 14-JUL-1999; 99US-0143624.
? 15-JUL-1999; 99US-0144005.
? 16-JUL-1999; 99US-0144085.
? 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

13-OCT-1999;	99US-0159293.	PR	05-MAR-1999;	99US-0123180.
13-OCT-1999;	99US-0159294.	PR	09-MAR-1999;	99US-0123548.
13-OCT-1999;	99US-0159295.	PR	23-MAR-1999;	99US-0125788.
14-OCT-1999;	99US-0159329.	PR	25-MAR-1999;	99US-0126264.
14-OCT-1999;	99US-0159330.	PR	29-MAR-1999;	99US-0126785.
14-OCT-1999;	99US-0159331.	PR	01-APR-1999;	99US-0127462.
14-OCT-1999;	99US-0159637.	PR	06-APR-1999;	99US-0128234.
14-OCT-1999;	99US-0159638.	PR	08-APR-1999;	99US-0128714.
18-OCT-1999;	99US-0159584.	PR	16-APR-1999;	99US-0129845.
21-OCT-1999;	99US-0160741.	PR	19-APR-1999;	99US-0130077.
21-OCT-1999;	99US-0160768.	PR	21-APR-1999;	99US-0130449.
21-OCT-1999;	99US-0160770.	PR	23-APR-1999;	99US-0130510.
21-OCT-1999;	99US-0160814.	PR	28-APR-1999;	99US-0130891.
21-OCT-1999;	99US-0160815.	PR	30-APR-1999;	99US-0131449.
22-OCT-1999;	99US-0160980.	PR	30-APR-1999;	99US-0132048.
22-OCT-1999;	99US-0160981.	PR	04-MAY-1999;	99US-0132407.
22-OCT-1999;	99US-0160989.	PR	05-MAY-1999;	99US-0132484.
25-OCT-1999;	99US-0161404.	PR	06-MAY-1999;	99US-0132485.
25-OCT-1999;	99US-0161405.	PR	06-MAY-1999;	99US-0132486.
25-OCT-1999;	99US-0161406.	PR	07-MAY-1999;	99US-0132487.
26-OCT-1999;	99US-0161359.	PR	11-MAY-1999;	99US-0132863.
26-OCT-1999;	99US-0161360.	PR	14-MAY-1999;	99US-0134256.
26-OCT-1999;	99US-0161361.	PR	14-MAY-1999;	99US-0134218.
28-OCT-1999;	99US-0161920.	PR	14-MAY-1999;	99US-0134219.
28-OCT-1999;	99US-0161992.	PR	14-MAY-1999;	99US-0134221.
28-OCT-1999;	99US-0161993.	PR	18-MAY-1999;	99US-0134370.
29-OCT-1999;	99US-0162142.	PR	18-MAY-1999;	99US-0134768.
ry Match 76.0%; Score 833.5; DB 21; Length 512;				
t Local Similarity 76.5%; Pred. No. 1.6e-82;				
ches 143; Conservative 29; Mismatches 14; Indels 1; Gaps 1;				
6 QDGKTEVIAWYMDSEEDQRLPHRPEKFEIPVDKLTGLVSWRLNPNWNCENLKR 65				
319 KDGREEVIAWYMDSEEDQRLPHRDPREFLSLDKLAELGVLWSRLDADNYETDEDLK 378				
66 IRARGYSYDICOVCEKLPNTYETKIKSFEEHLHTDEIRYCLSGSGYFVDRDQNDQW 125				
379 IRESRGYSYDMFCVCEKLPNTYEVKSPFEEHLHTDEIRYCVAGSGYFVDRDRNEAW 438				
126 IRTALKKGMIVLPAGMYHRETLDTONYIKAMRLFVCDPVMTYPNRPDHLPARKEFLAK 185				
439 IRVWVKKGMIVLPAGIYHRTVDSONYIKAMRLFVCEPVMTYPNRPDHLPARKEYIDN 498				
186 LLK-SEG 191				
499 FVKVNEG 505				
.T 9				
5299				
AAG45299 standard; Protein; 192 AA.				
AAG45299;				
18-OCT-2000 (first entry)				
Arabidopsis thaliana protein fragment SEQ ID NO: 56853.				
Protein identification; signal transduction pathway; metabolic pathway;				
hybridisation assay; genetic mapping; gene expression control; promoter;				
termination sequence.				
Arabidopsis thaliana.				
EP1033405-A2.				
06-SEP-2000.				
25-FEB-2000; 2000EP-0301439.				
25-FEB-1999; 99US-0121825.				

14-JUL-1999; 99US-014362A.
15-JUL-1999; 99US-014400S.
16-JUL-1999; 99US-014408S.
16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-014432S.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
19-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-014433S.
20-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144884.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145086.
21-JUL-1999; 99US-0145088.
22-JUL-1999; 99US-014508S.
22-JUL-1999; 99US-0145087.
22-JUL-1999; 99US-0145089.
22-JUL-1999; 99US-0145192.
23-JUL-1999; 99US-014514S.
23-JUL-1999; 99US-0145218.
23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145913.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
28-JUL-1999; 99US-0145951.
28-JUL-1999; 99US-0146386.
02-AUG-1999; 99US-0146388.
02-AUG-1999; 99US-0146389.
03-AUG-1999; 99US-0147038.
04-AUG-1999; 99US-0147204.
04-AUG-1999; 99US-0147302.
05-AUG-1999; 99US-0147192.
05-AUG-1999; 99US-0147260.
06-AUG-1999; 99US-0147303.
06-AUG-1999; 99US-0147416.
09-AUG-1999; 99US-0147493.
09-AUG-1999; 99US-014793S.
10-AUG-1999; 99US-0148171.
11-AUG-1999; 99US-0148319.
12-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-014856S.
13-AUG-1999; 99US-0148684.
16-AUG-1999; 99US-0149368.
17-AUG-1999; 99US-014917S.
18-AUG-1999; 99US-0149426.
20-AUG-1999; 99US-0149722.
20-AUG-1999; 99US-0149723.
20-AUG-1999; 99US-0149929.
23-AUG-1999; 99US-0149902.
23-AUG-1999; 99US-0149930.
25-AUG-1999; 99US-0150566.
26-AUG-1999; 99US-0150884.
27-AUG-1999; 99US-015106S.
27-AUG-1999; 99US-0151066.
27-AUG-1999; 99US-0151080.
30-AUG-1999; 99US-0151303.
31-AUG-1999; 99US-0151438.
01-SEP-1999; 99US-0151930.
07-SEP-1999; 99US-0152363.
10-SEP-1999; 99US-0153070.
13-SEP-1999; 99US-0153758.
15-SEP-1999; 99US-0154018.
16-SEP-1999; 99US-0154039.
20-SEP-1999; 99US-0154779.
22-SEP-1999; 99US-0155139.
23-SEP-1999; 99US-0155486.
24-SEP-1999; 99US-0155659.
28-SEP-1999; 99US-0156458.
29-SEP-1999; 99US-0156596.
04-OCT-1999; 99US-0157117.

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-015786S.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-015929S.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-016081S.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-016140S.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 75.5%; Score 828; DB 21; Length 192;

Best Local Similarity 76.5%; Pred. No. 1.6e-82;

Matches 140; Conservative 28; Mismatches 15; Indels 0; Gaps 0;

QY 1 MENEFDGKTEVTEAWYDDSEDORLPHRPEKPIPVDKLTGLGVTSWRLNPDWENC 60
Db 1 MGEAVKDGREVIQAWYDDSEDORLPHKOPKEVSLDKLAELGVLSWRLDADNYTD 60
QY 61 ENLKRIREARGSYVDICDVCPEKLPNETKIKSFEEHLHTDEEIRYCLESGYFDVRD 120
Db 61 EDLKKIRESRGYSYMBDFCEVCEPKLPNEYVKVKSFEHLHTDEEIRYCVAGTYFDVRD 120
QY 121 QNDQWIRIALKKGGMIVLPAGMYHRTLTQDNVYIKAMRLFVGDVPWTPYNNRPHDLPAK 180
Db 121 RNEAWIRVLVKGGMIVLPAGIYHRTVDSQDNVYIKAMRLFVGEVPWTPYNNRPHDLPAK 180
QY 181 EFL 183
Db 181 EYV 183

RESULT 10

AAU72532

ID AAU72532 standard; Protein; 559 AA.

XX AAU72532;

XX

DT 26-FEB-2002 (first entry)

XX Arabidopsis cell cycle protein CCP22.

DE Arabidopsis cell cycle protein CCP22.

XX Cell cycle protein; CCP; cell cycle regulation; herbicide;

KW plant growth regulator; plant development; abiotic stress; biotic stress;

KW nutrient deprivation; pathogen attack; crop yield; immunogen.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX WO200185946-A2.

PN

XX

-NOV-2001.
-MAY-2001; 2001WO-IB01307.
-MAY-2000; 2000US-204045P.
*ROP-) CROPDESIGN NV.
ize D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;
>1; 2002-062249/08.
-PSDB; AAS96322.
aw cell cycle protein and nucleic acid molecule encoding it useful for
regulating cell cycle progression in plants and for identifying
modulators which are useful as herbicides or plant growth regulators -
claim 34; Fig 23; 316pp; English.
he invention relates to a novel cell cycle protein (CCP) and the
olynucleotides encoding them. CCP is useful for identifying a compound
hich modulates the activity of the polypeptide and which binds to the
olypeptide and an anti-CCP antibody is useful for detecting the presence
f CCP in a sample. A CCP modulator is useful for modulating the cell
ycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,
maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.
CP nucleic acid and polypeptide molecules are useful as modulating
agents in regulating cell cycle progression in plants. CCP is useful to
reat disorders characterised by insufficient or excessive production of
CP protein or production of CCP protein forms which have decreased or
aberrant activity. Compounds that bind to or modulate the activity
of CCP polypeptide are useful as herbicides or plant growth regulators.
The polynucleotide is useful for modifying cell fate, plant development,
plant morphology, biochemistry and/or physiology, the length of the G1,
S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,
stimulation or enhancement of cell division, DNA replication, seed set,
seed size, seed development, tuber, fruit, leaf formation, shoot and root
initiation and/or development, nodule function, dwarfism in plants,
senescence, tolerance or resistance to stress. CCP, the polynucleotide
and the anti-CCP antibody are useful in agriculture to modulate the
protein levels or activity of a protein involved in the cell cycle due
to environmental conditions, including abiotic stress such as
cold, nutrient deprivation, heat, drought, salt stress, or biotic
stress such as pathogen attack, to modulate e.g. enhance crop yields,
and attenuate plant architecture, plant quality traits, plant
reproduction and seed development, endoreduplication in storage cells,
storage tissues and/or storage organs of plants or its parts. CCP is
useful as an immunogen to generate antibodies. CCP protein is useful to
screen for naturally occurring CCP substrates. The polynucleotide is
useful for expressing CCP protein, to detect CCP mRNA, or a genetic
lesion in a CCP gene and to modulate CCP activity. The present sequence
represents a CCP protein of the invention.
Sequence 559 AA;
ery Match 74.28; Score 814; DB 23; Length 559;
st Local Similarity 78.3%; Pred. No. 2.5e-80;
tches 137; Conservative 27; Mismatches 11; Indels 0; Gaps 0;
9 RTEVTEAWYMDSEEDQRLPHHREKPEFIPVDKLTGLGVISWRLNPWNENCENLKRI 68
376 REEVIQAWYMDSEEDQRLPHHREKPEFVSLDKLAEGLVSLWRDLADNYETDEELKIRE 435
69 ARGYSYVDICDVCPEKLPNYETKIKSFFEEHLHTDEEIRYCLGSGYFDVRDNDQWIRI 128
436 SRGYSYMDFCVCEVCPEKLPNYEVKVSFEEHLHTDEEIRYCVAGSYFDVRDNEAWIR 495
129 ALKKGGMIVLPAGMYHRTLDTDNYIKAMRLFVGPVWTPYNNRPHDHLPAKKEFL 183
496 LVKKGGMIVLPAGIYHRTVSDSNYIKAMRLFVGPVWTPYNNRPHDHLPAKKEV 550

JLT 11

AAE08931
ID AAE08931 standard; Protein; 196 AA.
XX
AC AAE08931;
XX
DT 15-NOV-2001 (first entry)
XX
DE Tomato submergence-induced early response protein SH2A homologue #1.
XX
KW Tomato; SH2A-like protein; submergence induced early response gene;
KW anaerobic response; water logging tolerance; hypoxic environment;
KW horticultural; agricultural; medical; fermentation;
KW cell culture industry; genetic construct; transgenic plant.
XX
OS Lycopersicon esculentum.
XX
PN WO200161020-A2.
XX
PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001WO-IB00745.
XX
PR 18-FEB-2000; 2000US-0183572.
XX
PA (CROP-) CROPDESIGN NV.
XX
PI Sauter MM, Lorbiecke R;
DR WPI: 2001-529915/58.
DR N-PSDB; AAD15649.
XX
New genetic constructs, comprising the SH2A or SH2A-like genes, useful
for altering the growth and adaptation of plants under hypoxic
conditions, e.g. improving water logging tolerance or regulating
anaerobic response in plants -
XX
PS Claim 32; Page 95-96; 107pp; English.
XX
The present invention relates to genetic constructs, which comprise a
nucleotide sequence for an SH2A or SH2A-like gene, or an SH2A or SH2A-
like protein gene promoter, where the SH2A or SH2A-like gene is operably
linked to a promoter sequence that directs expression of the nucleotide
sequence and the SH2A or SH2A-like protein gene promoter is operably
linked to a heterologous coding sequence. The SH2A gene, initially
identified in rice plants, is a submergence-induced early response gene
(such as those encountered during submersion). In addition, the SH2A gene
product has also been found to belong to a family of highly conserved
proteins which occur ubiquitously in eukaryotes. The genetic constructs
and methods are useful for altering the growth and adaptation of plants
under hypoxic conditions. In particular, these are useful for improving
water logging tolerance and regulating anaerobic response in plants.
These are particularly useful in horticultural, agricultural, medical,
fermentation and cell culture industries. The present sequence is a
tomato SH2A-like protein.
XX
SQ Sequence 196 AA;
Query Match 74.0%; Score 812; DB 22; Length 196;
Best Local Similarity 73.2%; Pred. No. 9.5e-81;
Matches 139; Conservative 28; Mismatches 23; Indels 0; Gaps 0;
QY 7 DCKTEVTEAWYMDSEEDQRLPHHREKPEFIPVDKLTGLGVISWRLNPWNENCENLKRI 66
DB 3 DREVDIQAWYMDNDDEDQRLPHHREKPEFVSLDKLAEGLVSLWRDLADNYETDEELKKI 62
QY 67 REARGYSYVDICDVCPEKLPNYETKIKSFFEEHLHTDEEIRYCLGSGYFDVRDNDQWIRI 126
DB 63 REDRGYSYMDFCVCEVCPEKLPNYEKKNFEEHLHTDEEIRYAVAGSYFDVRDNEWSWI 122
QY 127 RYALKKGGMIVLPAGMYHRTLDTDNYIKAMRLFVGPVWTPYNNRPHDHLPAKKEFLAKL 186
DB 123 RYVKKKGGMIVLPAGIYHRTVSDSNYIKAMRLFVGPVWTPYNNRPHDHLPAKKEVYETF 182

Y 187 LKSEGENQAV 196
: : : : :
b 183 VNADGAGRAV 192

3SULT 12
AC27918
> AAG27918 standard; Protein; 198 AA.
{
AAG27918;

17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 32943.

Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

Arabidopsis thaliana.

EP1033405-A2.

06-SEP-2000.

25-FEB-2000; 2000EP-0301439.

25-FEB-1999; 99US-0121825.

05-MAR-1999; 99US-0123180.

09-MAR-1999; 99US-0123548.

23-MAR-1999; 99US-0125788.

25-MAR-1999; 99US-0126264.

29-MAR-1999; 99US-0126785.

01-APR-1999; 99US-0127462.

06-APR-1999; 99US-0128234.

08-APR-1999; 99US-0128714.

16-APR-1999; 99US-0129845.

19-APR-1999; 99US-0130077.

21-APR-1999; 99US-0130449.

23-APR-1999; 99US-0130510.

28-APR-1999; 99US-0130891.

30-APR-1999; 99US-0131449.

04-MAY-1999; 99US-0132407.

05-MAY-1999; 99US-0132484.

06-MAY-1999; 99US-0132485.

07-MAY-1999; 99US-0132486.

11-MAY-1999; 99US-0132487.

14-MAY-1999; 99US-0132653.

14-MAY-1999; 99US-0134218.

14-MAY-1999; 99US-0134219.

14-MAY-1999; 99US-0134221.

18-MAY-1999; 99US-0134370.

19-MAY-1999; 99US-0134768.

20-MAY-1999; 99US-0134941.

21-MAY-1999; 99US-0135124.

24-MAY-1999; 99US-0135553.

25-MAY-1999; 99US-0135629.

27-MAY-1999; 99US-0136021.

28-MAY-1999; 99US-0136392.

01-JUN-1999; 99US-0136782.

03-JUN-1999; 99US-0137222.

04-JUN-1999; 99US-0137528.

07-JUN-1999; 99US-0137724.

08-JUN-1999; 99US-0138094.

10-JUN-1999; 99US-0138540.

10-JUN-1999; 99US-0138847.

14-JUN-1999; 99US-0139119.

16-JUN-1999; 99US-0139452.

16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.

99US-0138540.
10-JUN-1999;
99US-0138847.
10-JUN-1999;
99US-0139119.
14-JUN-1999;
99US-0139452.
16-JUN-1999;
99US-0139453.
16-JUN-1999;
99US-0139492.
17-JUN-1999;
99US-0139454.
18-JUN-1999;
99US-0139455.
18-JUN-1999;
99US-0139456.
18-JUN-1999;
99US-0139457.
18-JUN-1999;
99US-0139458.
18-JUN-1999;
99US-0139459.
18-JUN-1999;
99US-0139460.
18-JUN-1999;
99US-0139461.
18-JUN-1999;
99US-0139462.
18-JUN-1999;
99US-0139463.
18-JUN-1999;
99US-0139750.
18-JUN-1999;
99US-0139763.
21-JUN-1999;
99US-0139817.
21-JUN-1999;
99US-0139899.
23-JUN-1999;
99US-0140353.
23-JUN-1999;
99US-0140354.
23-JUN-1999;
99US-0140695.
24-JUN-1999;
99US-0140823.
28-JUN-1999;
99US-0140991.
29-JUN-1999;
99US-014287.
30-JUN-1999;
99US-0141287.
01-JUL-1999;
99US-0141842.
01-JUL-1999;
99US-0142154.
02-JUL-1999;
99US-0142055.
06-JUL-1999;
99US-0142390.
08-JUL-1999;
99US-0142803.
09-JUL-1999;
99US-0142920.
12-JUL-1999;
99US-0143977.
13-JUL-1999;
99US-0143542.
14-JUL-1999;
99US-0143624.
15-JUL-1999;
99US-0144005.
16-JUL-1999;
99US-0144085.
16-JUL-1999;
99US-0144086.
19-JUL-1999;
99US-0144325.
19-JUL-1999;
99US-0144331.
19-JUL-1999;
99US-0144332.
19-JUL-1999;
99US-0144333.
19-JUL-1999;
99US-0144334.
19-JUL-1999;
99US-0144335.
20-JUL-1999;
99US-0144352.
20-JUL-1999;
99US-0144632.
20-JUL-1999;
99US-0144884.
21-JUL-1999;
99US-0144814.
21-JUL-1999;
99US-0145086.
21-JUL-1999;
99US-0145088.
22-JUL-1999;
99US-0145085.
22-JUL-1999;
99US-0145087.
22-JUL-1999;
99US-0145089.
22-JUL-1999;
99US-0145192.
23-JUL-1999;
99US-0145145.
23-JUL-1999;
99US-0145218.
23-JUL-1999;
99US-0145224.
26-JUL-1999;
99US-0145276.
27-JUL-1999;
99US-0145913.
27-JUL-1999;
99US-0145918.
27-JUL-1999;
99US-0145919.
28-JUL-1999;
99US-0145951.
02-AUG-1999;
99US-0146386.
02-AUG-1999;
99US-0146388.
02-AUG-1999;
99US-0146389.
03-AUG-1999;
99US-0147036.
04-AUG-1999;
99US-0147204.
04-AUG-1999;
99US-0147302.
05-AUG-1999;
99US-0147192.
05-AUG-1999;
99US-0147260.
06-AUG-1999;
99US-0147303.
06-AUG-1999;
99US-0147416.
09-AUG-1999;
99US-0147493.

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161484.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.


```
05-AUG-1999; 99US-0147192.
05-AUG-1999; 99US-0147260.
06-AUG-1999; 99US-0147303.
06-AUG-1999; 99US-0147416.
09-AUG-1999; 99US-0147493.
09-AUG-1999; 99US-0147935.
10-AUG-1999; 99US-0148171.
11-AUG-1999; 99US-0148319.
12-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-0148565.
13-AUG-1999; 99US-0148684.
16-AUG-1999; 99US-0149368.
17-AUG-1999; 99US-0149175.
18-AUG-1999; 99US-0149426.
20-AUG-1999; 99US-0149722.
20-AUG-1999; 99US-0149723.
20-AUG-1999; 99US-0149929.
23-AUG-1999; 99US-0149902.
23-AUG-1999; 99US-0149930.
25-AUG-1999; 99US-0150566.
26-AUG-1999; 99US-0150884.
27-AUG-1999; 99US-0151065.
27-AUG-1999; 99US-0151066.
27-AUG-1999; 99US-0151080.
30-AUG-1999; 99US-0151303.
31-AUG-1999; 99US-0151438.
01-SEP-1999; 99US-0151930.
07-SEP-1999; 99US-0152363.
10-SEP-1999; 99US-0153070.
13-SEP-1999; 99US-0153758.
15-SEP-1999; 99US-0154018.
16-SEP-1999; 99US-0154039.
20-SEP-1999; 99US-0154779.
22-SEP-1999; 99US-0155139.
23-SEP-1999; 99US-0155486.
24-SEP-1999; 99US-0155659.
28-SEP-1999; 99US-0156458.
29-SEP-1999; 99US-0156596.
04-OCT-1999; 99US-0157117.
05-OCT-1999; 99US-0157753.
06-OCT-1999; 99US-0157865.
07-OCT-1999; 99US-0158029.
08-OCT-1999; 99US-0158232.
12-OCT-1999; 99US-0158369.
13-OCT-1999; 99US-0159293.
13-OCT-1999; 99US-0159294.
13-OCT-1999; 99US-0159295.
14-OCT-1999; 99US-0159329.
14-OCT-1999; 99US-0159330.
14-OCT-1999; 99US-0159331.
14-OCT-1999; 99US-0159637.
14-OCT-1999; 99US-0159638.
18-OCT-1999; 99US-0159584.
21-OCT-1999; 99US-0160741.
21-OCT-1999; 99US-0160767.
21-OCT-1999; 99US-0160768.
21-OCT-1999; 99US-0160770.
21-OCT-1999; 99US-0160814.
21-OCT-1999; 99US-0160815.
22-OCT-1999; 99US-0160980.
22-OCT-1999; 99US-0160981.
22-OCT-1999; 99US-0160989.
25-OCT-1999; 99US-0161404.
25-OCT-1999; 99US-0161405.
26-OCT-1999; 99US-0161406.
26-OCT-1999; 99US-0161359.
26-OCT-1999; 99US-0161360.
26-OCT-1999; 99US-0161361.
28-OCT-1999; 99US-0161920.
28-OCT-1999; 99US-0161922.
28-OCT-1999; 99US-0161993.
29-OCT-1999; 99US-0162142.
29-OCT-1999; 99US-0162142.

Query Match 71.3%; Score 782.5; DB 21; Length 182;
Best Local Similarity 77.1%; Pred. No. 1.5e-77;
Matches 135; Conservative 26; Mismatches 13; Indels 1; Gaps 1;

QY 18 MDDSEEDQRLPHHREPKEFIPVDKLTGLGVISWRNLNPDNWCENLKRIRARGSYVDI 77
Db 1 MDDSEEDQRLPHHKDFKFLSLDKLAELGVLSWRDLADNYETDEDLKIRSRGYSYMF 60
QY 78 CDVCEKLPNYETKIKSFEEHLHTDEIRYCLEGSYFVDRDNDOWIRIALKKGGMIV 137
Db 61 CEVCEKLPNYEVKVSFEEHLHTDEIRYCVAGSYFVDRDNEAIRVWVKKGGMIV 120
QY 138 LPAGMYHRTDLDNYIKAMRLFVGDPVWTPYNRPHDHPARKEFLAKLLK-SEG 191
Db 121 LPAGIVHRTVDSNDYIKAMRLFVGEPVWTPYNRPHDHPARKEYIDNFVKVNEG 175

RESULT 15
AAG45306
ID AAG45306 standard; Protein; 182 AA.
XX
AC AAG45306;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 56862.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
```

8-MAY-1999; 99US-0136782.
1-JUN-1999; 99US-0137222.
13-JUN-1999; 99US-0137528.
4-JUN-1999; 99US-0137502.
17-JUN-1999; 99US-0137724.
18-JUN-1999; 99US-0138094.
0-JUN-1999; 99US-0138540.
0-JUN-1999; 99US-0138847.
4-JUN-1999; 99US-0139119.
6-JUN-1999; 99US-0139452.
6-JUN-1999; 99US-0139453.
7-JUN-1999; 99US-0139492.
8-JUN-1999; 99US-0139454.
18-JUN-1999; 99US-0139455.
18-JUN-1999; 99US-0139456.
18-JUN-1999; 99US-0139457.
18-JUN-1999; 99US-0139458.
18-JUN-1999; 99US-0139459.
18-JUN-1999; 99US-0139460.
18-JUN-1999; 99US-0139461.
18-JUN-1999; 99US-0139462.
18-JUN-1999; 99US-0139463.
18-JUN-1999; 99US-0139750.
18-JUN-1999; 99US-0139763.
21-JUN-1999; 99US-0139817.
22-JUN-1999; 99US-0139899.
23-JUN-1999; 99US-0140353.
23-JUN-1999; 99US-0140354.
24-JUN-1999; 99US-0140695.
28-JUN-1999; 99US-0140823.
28-JUN-1999; 99US-0140991.
30-JUN-1999; 99US-0141287.
01-JUL-1999; 99US-0141842.
01-JUL-1999; 99US-0142154.
02-JUL-1999; 99US-0142055.
06-JUL-1999; 99US-0142390.
08-JUL-1999; 99US-0142803.
09-JUL-1999; 99US-0142920.
12-JUL-1999; 99US-0142977.
13-JUL-1999; 99US-0143542.
14-JUL-1999; 99US-0143624.
15-JUL-1999; 99US-0144005.
16-JUL-1999; 99US-0144085.
16-JUL-1999; 99US-0144086.
19-JUL-1999; 99US-0144325.
19-JUL-1999; 99US-0144331.
19-JUL-1999; 99US-0144332.
19-JUL-1999; 99US-0144333.
19-JUL-1999; 99US-0144334.
19-JUL-1999; 99US-0144335.
20-JUL-1999; 99US-0144352.
20-JUL-1999; 99US-0144632.
20-JUL-1999; 99US-0144884.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0144814.
21-JUL-1999; 99US-0145086.
21-JUL-1999; 99US-0145088.
22-JUL-1999; 99US-0145085.
22-JUL-1999; 99US-0145087.
22-JUL-1999; 99US-0145089.
22-JUL-1999; 99US-0145192.
23-JUL-1999; 99US-0145145.
23-JUL-1999; 99US-0145218.
23-JUL-1999; 99US-0145224.
26-JUL-1999; 99US-0145276.
27-JUL-1999; 99US-0145913.
27-JUL-1999; 99US-0145918.
27-JUL-1999; 99US-0145919.
28-JUL-1999; 99US-0145951.
02-AUG-1999; 99US-0146386.
02-AUG-1999; 99US-0146388.
02-AUG-1999; 99US-0146389.
03-AUG-1999; 99US-0147038.
04-AUG-1999; 99US-0147204.
04-AUG-1999; 99US-0147302.
05-AUG-1999; 99US-0147192.
05-AUG-1999; 99US-0147260.
06-AUG-1999; 99US-0147303.
06-AUG-1999; 99US-0147416.
09-AUG-1999; 99US-0147493.
09-AUG-1999; 99US-0147935.
09-AUG-1999; 99US-0148171.
10-AUG-1999; 99US-0148319.
11-AUG-1999; 99US-0148341.
13-AUG-1999; 99US-0148565.
13-AUG-1999; 99US-0148684.
16-AUG-1999; 99US-0149368.
17-AUG-1999; 99US-0149175.
18-AUG-1999; 99US-0149426.
20-AUG-1999; 99US-0149722.
20-AUG-1999; 99US-0149723.
20-AUG-1999; 99US-0149929.
23-AUG-1999; 99US-0149902.
23-AUG-1999; 99US-0149930.
25-AUG-1999; 99US-0150566.
26-AUG-1999; 99US-0150884.
27-AUG-1999; 99US-0151065.
27-AUG-1999; 99US-0151066.
27-AUG-1999; 99US-0151080.
30-AUG-1999; 99US-0151303.
31-AUG-1999; 99US-0151438.
01-SEP-1999; 99US-0151930.
07-SEP-1999; 99US-0152363.
10-SEP-1999; 99US-0153070.
13-SEP-1999; 99US-0153758.
15-SEP-1999; 99US-0154018.
16-SEP-1999; 99US-0154039.
20-SEP-1999; 99US-0154779.
22-SEP-1999; 99US-0155139.
23-SEP-1999; 99US-0155486.
24-SEP-1999; 99US-0155659.
28-SEP-1999; 99US-0156458.
29-SEP-1999; 99US-0156596.
04-OCT-1999; 99US-0157117.
05-OCT-1999; 99US-0157753.
06-OCT-1999; 99US-0157865.
07-OCT-1999; 99US-0158029.
08-OCT-1999; 99US-0158232.
12-OCT-1999; 99US-0158369.
13-OCT-1999; 99US-0159293.
13-OCT-1999; 99US-0159294.
13-OCT-1999; 99US-0159295.
14-OCT-1999; 99US-0159329.
14-OCT-1999; 99US-0159330.
14-OCT-1999; 99US-0159331.
14-OCT-1999; 99US-0159637.
14-OCT-1999; 99US-0159638.
18-OCT-1999; 99US-0159584.
21-OCT-1999; 99US-0160741.
21-OCT-1999; 99US-0160767.
21-OCT-1999; 99US-0160768.
21-OCT-1999; 99US-0160770.
21-OCT-1999; 99US-0160814.
21-OCT-1999; 99US-0160815.
22-OCT-1999; 99US-0160980.
22-OCT-1999; 99US-0160981.
22-OCT-1999; 99US-0160989.
25-OCT-1999; 99US-0161404.
25-OCT-1999; 99US-0161405.
25-OCT-1999; 99US-0161406.
26-OCT-1999; 99US-0161359.
26-OCT-1999; 99US-0161360.
28-OCT-1999; 99US-0161361.
28-OCT-1999; 99US-0161920.
28-OCT-1999; 99US-0161992.
28-OCT-1999; 99US-0161993.
29-OCT-1999; 99US-0162142.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

n - protein search, using sw model

November 25, 2002, 10:08:20 : Search time 15 seconds
(without alignments)
390.344 Million cell updates/sec

score: 1097

US-09-785-738A-2 : REFLLAKLLKSEGENQAVGEF 199

table: BLOSUM62

Gapop 10.0 , Gapext 0.5

i: 262574 seqs, 29422922 residues

umber of hits satisfying chosen parameters: 262574

DB seq length: 0

DB seq length: 2000000000

ocessing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

se : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A-COMB.pap:*
- 2: /cgn2_6/ptodata/1/iaa/5B-COMB.pap:*
- 3: /cgn2_6/ptodata/1/iaa/6A-COMB.pap:*
- 4: /cgn2_6/ptodata/1/iaa/6B-COMB.pap:*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pap:*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

lt	Query	%	Score	Match	Length	DB	ID	Description
1	586	53.4	179	2	US-08-933-750C-28			Sequence 28, Appl
2	586	53.4	179	4	US-09-234-613-28			Sequence 28, Appl
3	439.5	40.1	163	4	US-09-348-265-2			Sequence 2, Appl
4	94.5	8.6	749	4	US-09-562-737-96			Sequence 96, Appl
5	84	7.7	323	4	US-09-237-543-6			Sequence 6, Appl
6	84	7.7	323	4	US-09-644-450-6			Sequence 26, Appl
7	81	7.4	313	2	US-08-463-081B-26			Sequence 26, Appl
8	81	7.4	313	2	US-08-461-379A-26			Sequence 26, Appl
9	81	7.4	313	2	US-08-462-390B-26			Sequence 26, Appl
10	81	7.4	313	3	US-08-463-074B-26			Sequence 26, Appl
11	81	7.4	313	3	US-08-465-585C-26			Sequence 26, Appl
12	81	7.4	313	3	US-08-652-446-26			Sequence 26, Appl
13	81	7.4	313	4	US-09-237-543-8			Sequence 8, Appl
14	81	7.4	313	4	US-09-237-543-9			Sequence 9, Appl
15	81	7.4	313	4	US-09-644-450-8			Sequence 8, Appl
16	81	7.4	313	4	US-09-644-450-9			Sequence 9, Appl
17	81	7.4	876	4	US-09-134-001C-3943			Sequence 3943, Ap
18	77.5	7.1	749	4	US-09-562-737-91			Sequence 91, Appl
19	77	7.0	257	2	US-07-857-224B-41			Sequence 41, Appl
20	77	7.0	313	4	US-09-237-543-7			Sequence 7, Appl
21	77	7.0	313	4	US-09-644-450-7			Sequence 7, Appl
22	77	7.0	318	4	US-09-134-001C-3603			Sequence 3603, Ap
23	76	6.9	571	1	US-07-955-905A-25			Sequence 25, Appl
24	75.5	6.9	749	4	US-09-562-737-99			Sequence 99, Appl
25	75.5	6.9	1138	4	US-09-580-101A-8			Sequence 8, Appl
26	75.5	6.9	1151	4	US-09-590-101A-10			Sequence 10, Appl
27	74.5	6.8	540	1	US-08-249-013-2			Sequence 2, Appl

28	74.5	6.8	540	2	US-08-886-863-2			Sequence 2, Appl
29	74.5	6.8	540	4	US-09-175-229-2			Sequence 2, Appl
30	74.5	6.8	540	5	PCT-US95-06764-2			Sequence 2, Appl
31	74	6.7	450	4	US-09-306-593-13			Sequence 13, Appl
32	72	6.6	305	4	US-09-134-001C-5632			Sequence 5632, Ap
33	71.5	6.5	345	2	US-08-446-345-40			Sequence 40, Appl
34	71	6.5	232	4	US-09-134-001C-3649			Sequence 3649, Ap
35	71	6.5	348	4	US-09-671-950-10			Sequence 10, Appl
36	71	6.5	348	4	US-09-671-950-14			Sequence 14, Appl
37	71	6.5	765	4	US-08-737-109-11			Sequence 11, Appl
38	70.5	6.4	284	2	US-08-766-439-44			Sequence 44, Appl
39	70.5	6.4	288	3	US-09-100-804-14			Sequence 14, Appl
40	70.5	6.4	706	4	US-09-134-001C-4908			Sequence 4908, Ap
41	70	6.4	158	4	US-09-042-071-20			Sequence 20, Appl
42	70	6.4	348	4	US-09-671-950-2			Sequence 2, Appl
43	70	6.4	348	4	US-09-671-950-4			Sequence 4, Appl
44	70	6.4	348	4	US-09-671-950-6			Sequence 6, Appl
45	70	6.4	348	4	US-09-671-950-12			Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-933-750C-28
; Sequence 28, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 179 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LNOBNOT03

CLONE: 1577239
US-08-933-750C-28

Query Match 53.4%; Score 586; DB 2; Length 179;
Best Local Similarity 56.7%; Pred. No. 7.8e-59;
Matches 101; Conservative 34; Mismatches 43; Indels 0; Gaps 0;
QY 12 VIEAWYMDSEEDQRLPHHREPEKIPVDKLTGLVSWRLNPNWNCENLKRIEARG 71
DB 1 MVQAWYMDADPGDPQPHRPDPGRPVGLEQLRLGLVLYWKLDAKYENDEPELEKIRRN 60
Y 72 YSVVDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLGSGYFDVRDQNDQWIRIA 131
DB 61 YSWMDIITICKDKLPNYEEKIKMFEEHLHLDDEIRYILDSGYFDVRDQNDQWIRI 120
Y 132 KGMIVLPAGMYHFTLTDNYIKAMRLFVGDPVWTPYNNRPHDLPARKEFLAKLKS 189
DB 121 KGMVTLPGIYHRTVDENKNTKAMRLFVGEPVWTAYNRPADHFEARGQYVRF 178

RESULT 2
US-09-234-613-28
Sequence 28, Application US/09234613
Patent No. 6132973

GENERAL INFORMATION:

APPLICANT: Lal Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purvi
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/234,613
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750
FILING DATE: September 23, 1997
TITNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:

RMATION FOR SEQ ID NO: 28:

QUENCE CHARACTERISTICS:

LENGTH: 179 amino acids

TYPE: amino acid

TRANSDUCEDNESS: single

OPOLGY: linear

EDATE SOURCE:

LIBRARY: LNC000703

CLONE: 1577239

4-613-28

Query Match 53.4%; Score 586; DB 4; Length 179;
Best Local Similarity 56.7%; Pred. No. 7.8e-59;
Matches 101; Conservative 34; Mismatches 43; Indels 0; Gaps 0;
QY 12 VIEAWYMDSEEDQRLPHHREPEKIPVDKLTGLVSWRLNPNWNCENLKRIEARG 71
DB 1 MVQAWYMDADPGDPQPHRPDPGRPVGLEQLRLGLVLYWKLDAKYENDEPELEKIRRN 60
QY 72 YSVVDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLGSGYFDVRDQNDQWIRIA 131
DB 61 YSWMDIITICKDKLPNYEEKIKMFEEHLHLDDEIRYILDSGYFDVRDQNDQWIRI 120
QY 132 KGMIVLPAGMYHFTLTDNYIKAMRLFVGDPVWTPYNNRPHDLPARKEFLAKLKS 189
DB 121 KGMVTLPGIYHRTVDENKNTKAMRLFVGEPVWTAYNRPADHFEARGQYVRF 178

RESULT 3

US-09-348-265-2
Sequence 2, Application US/09348265
Patent No. 6444800

GENERAL INFORMATION:

APPLICANT: KIKUCHI, Kokichi
APPLICANT: SATO, No. 6444800iyuki
APPLICANT: TORIGOE, Toshihiko
APPLICANT: SAHARA, Hiroeki
APPLICANT: SUZUKI, Manabu
APPLICANT: HAMURO, Junji
TITLE OF INVENTION: Human Gastric Cancer Antigen Gene and Gastric
Cancer Antigen Protein
FILE REFERENCE: OP871
CURRENT APPLICATION NUMBER: US/09/348,265
CURRENT FILING DATE: 1999-07-07
EARLIER APPLICATION NUMBER: JP 10-197852
EARLIER FILING DATE: 1998-07-13
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2

LENGTH: 163

TYPE: PRT

ORGANISM: Homo sapiens

US-09-348-265-2

Query Match 40.1%; Score 439.5; DB 4; Length 163;
Best Local Similarity 52.9%; Pred. No. 3.2e-42;
Matches 81; Conservative 27; Mismatches 42; Indels 3; Gaps 2;

QY 12 VIEAWYMD--DSEEDQRLPHHREPEKIPVDKLTGLVSWRLNPNWNCENLKRIE 69
DB 1 MVQAWYMGRRPGRPAATPPRPAQW-GLSQLRLGLVLYWKLDAKYENDEPELEKIR 59
QY 70 RGSYVDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLGSGYFDVRDQNDQWIRIA 129
DB 60 RNSWMDIITICKDKLPNYEEKIKMFEEHLHLDDEIRYILDSGYFDVRDQNDQWIRI 119
QY 130 LKGMIVLPAGMYHFTLTDNYIKAMRLFVG 162
DB 120 MERGDMVTLPGIYHRTVDENKNTKAMRLVCG 152

RESULT 4

US-09-562-737-96
Sequence 96, Application US/09562737
Patent No. 6428967

GENERAL INFORMATION:

APPLICANT: Herz, Joachim
APPLICANT: Gotthardt, Michael
TITLE OF INVENTION: LDL Receptor Signaling Pathways
FILE REFERENCE: UTSW0708
CURRENT APPLICATION NUMBER: US/09/562,737
CURRENT FILING DATE: 2000-05-01
NUMBER OF SEQ ID NOS: 96

```

) NO 96
:TH: 749
:; PRT
ANIM: Artificial Sequence
PURE:
ER INFORMATION: Description of Artificial Sequence: Synthetic
ER INFORMATION: Sequence
62-737-96

Match      8.6%; Score 94.5; DB 4; Length 749;
Local Similarity 23.2%; Pred. NO. 0.051;
ies 42; Conservative 25; Mismatches 65; Indels 49; Gaps

9 KTEVIEAWYDDSEEDQRLPHHREPK---EFTPVDKLTGLGVISRWLNDNW---ENC 60
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
5 KAESVGSGLMDRRVRPGVPVHQEPSEDEMF-PLGVYPEGELAAALRPESHAPSEQEC 63
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
61 ENLKRIREARGYSYVDICDVCPEKPLNYETKISFEEHLHDFE-----IRYCLEGSY 115
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
64 HNIS-----PDGSSSDKVN--TSEEDLDEGLPEEEMITYIRYCPNDDSY 110
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
116 FDVROQN-DQWIRIALKKGMLVLPAGMYHRTLDNYIKAMRLFGCDPWPVTPNRPDH 174
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :
111 LEGHQNGEYL-----AHSRHPVDDECSAVE-----EWTDSATPHP 149
||| : : : ||| : : : ||| : : : ||| : : : ||| : : :

175 H 175
150 H 150

```

```

T 5
-237-543-6
uence 6, Application US/09237543A
ent No. 6143540
ERAL INFORMATION:
PLICANT: Kapeller, Rosana
TITLE OF INVENTION: NOVEL MOLECULES OF THE HKID-1-RELATED PROTEIN FAMILY
FILE REFERENCE: 035800/175631
URRENT APPLICATION NUMBER: US/09/237,543A
URRENT FILING DATE: 1999-01-26
UMBER OF SEQ ID NOS: 11
FTWARE: PatentIn Ver. 2.0
Q ID NO 6
LENGTH: 323
TYPE: PRT
ORGANISM: Xenopus laevis
)9-237-543-6

ery Match 7.7%; Score 84; DB 4; Length 323;
est Local Similarity 24.2%; Pred. No. 0.24;
atches 37; Conservative 27; Mismatches 57; Indels 32; Gaps
33 PKFEIPIVDKLT-----LGVISWRLNPDNW---ENCNLRKTRKARGYSYVDICDVCPE 83
88 PLEIVLLKKVPTAFRGVINLLDWERPDALIVMERPEPVK-----DLFDVITE 136
84 KLPNVETKIKSFEEHLHTDEIRCYCLE-SSGYFDYDRQNDQWIRALKKG--GMIVLPA 140
137 KGPLDEDTARGFRQVL---EAVRHCYNCGVVHRDKDEN---LLVDTRNGELKLDIFGS 190
141 GMYHREFTLDTDNIKAMRLEFVGDPVWTPYNRPH 173
191 GALLKDTVTD--FDGRTVY--SPPEWRYHRH 220

```

SULT 6
-09-644-450-6
Sequence 6, Application US/09644450
Patent No. 6383791
GENERAL INFORMATION:
APPLICANT: Kapeller, Rosana
TITLE OF INVENTION: NOVEL MOLECULES OF THE HKID-1-RELATED PROTEIN FAMILY

```

; TITLE OF INVENTION:  AND USES THEREOF
; FILE REFERENCE:  035600/175631
; CURRENT APPLICATION NUMBER:  US/09/644,450
; CURRENT FILING DATE:  2000-08-23
; NUMBER OF SEQ ID NOS:  11
; SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH:  323
; TYPE:  PRT
; ORGANISM:  Xenopus laevis
;
US-09-644-450-6

Query Match      7.7%  Score 84;  DB 4;  Length 323;
Best Local Similarity 24.2%;  Pred. No. 0.24;
Matches 37;  Conservative 27;  Mismatches 57;  Indels 32;  Gaps 9;

Qy      33  PKFEIPVDKLTE-----LGVISRWLNPNW-----ENCENLKRIREARGSYVVDICDVCPE 83
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      88  PLEIVLLKKVPTAFRGVINLLDWERDPAFLVMEPEPVK-----DLFDYTE 136

Qy      84  KLPNYETKIKSFEEHLTDEIRYCLE-GSGYFDVDRDNDOWIRIAKKG--GMIVLPA 140
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      137  KGPLDEDTARGFFROVL---GAVRHCTNCGVHVHIDKEN---LLVDFRNGELKLIDFGS 190

Qy      141  GMVHRETLTDONYIKAMBLFVGDFVPTYNRPH 173
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      191  GALLKQTVTVD--FDGTRVY-SPPEWVRYIRYH 220

```

RESULT 7
US-08-463-081B-26
Sequence 26, Application US/08463081B
Patent No. 5871960
Patent No. 5871960 5837487
GENERAL INFORMATION:
APPLICANT: Smith, Kendall A. & Beadling, Carol
TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,
Vector and Transformed Cell Thereof, and Expression Thereof
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
STREET: 444 South Flower St. - Suite 1900
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463.081B
FILING DATE: 5-JUN-1995
PRIOR APPLICATION NUMBER: US 08/104,736
FILING DATE: 10-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/796,066
FILING DATE: 20-NOV-91
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Amzel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38150 (DART-060)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 622-7700
TELEFAX: (213) 489-4210
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 313 amino acids
TYPE: amino acid
STRANDEDNESS: n.a.

Db	85	RVPMEVLLKKVSSGSGVIRLLDWFPERDPSFVLLIERPEVQ	-----DLDFDI 133
QY	82	PEKLPNYETKIKSFEEHLHTDEEIRYCLE-GSGYFDVRDQNDQWIRALKKG--GMIVL 138	
Db	134	TERGALOELARSEFFWQVL---EAVRHCHNCVGLHRDKDEN---ILIDLNRGELKLIDF 187	
QY	139	PAGMYHRTLTDTDNVYIKAMRLFVGDVPWTPYNRPH 173	
Db	188	GSGALLKDTVTYD--FDGTRVY-SPPEWIRYHRYH 219	
<p>RESULT 9</p> <p>US-08-462-390B-26</p> <p>Sequence 26, Application US/08462390B</p> <p>Patent No. 5882894</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Smith, K. A., & Beadling, C.</p> <p>TITLE OF INVENTION: Nucleic Acids Encoding CR8 Polypeptide, Vector and</p> <p>TITLE OF INVENTION: Transformed Cell Thereof, and Expression Thereof</p> <p>NUMBER OF SEQUENCES: 35</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: Ratner & Prestia (B) STREET:One Westlakes-Berwyn</p> <p>CITY: Valley Forge</p> <p>STATE: Pennsylvania</p> <p>COUNTRY: USA</p> <p>ZIP: 19482</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: PatentIn Release #1.0, Version #1.25</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/462,390B</p> <p>FILING DATE: 5-JUNE-1995</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: USSN 08/330,108</p> <p>FILING DATE: 27-OCT-1994</p> <p>APPLICATION NUMBER: USSN 08/104,736</p> <p>FILING DATE: 10-AUG-1993</p> <p>APPLICATION NUMBER: USSN 07/796,066</p> <p>FILING DATE: 20-NOV-91</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Viviana Amzel, Ph. D.</p> <p>REGISTRATION NUMBER: 30,930</p> <p>REFERENCE/DOCKET NUMBER: DART-040</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: (610)407-0700</p> <p>TELEFAX: (610)407-0701</p> <p>INFORMATION FOR SEQ ID NO: 26:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 313 amino acids</p> <p>TYPE: peptide</p> <p>STRANDEDNESS: n.a.</p> <p>TOPOLOGY: n.a.</p> <p>MOLECULE TYPE: peptide</p> <p>US-08-462-390B-26</p>			
<p>Query Match 7.4%; Score 81; DB 2; Length 313;</p> <p>Best Local Similarity 25.2%; Pred. No. 0.5;</p> <p>Matches 39; Conservative 29; Mismatches 55; Indels 32; Gaps</p>			
QY	31	REPKEFIPVDKLT-----LGVISWRLLNPDNW---ENCENLKRREARGSYVDICDVC 81	
Db	85	RVPMEVLLKKVSSGSGVIRLLDWFPERDPSFVLLIERPEVQ	-----DLDFDI 133
QY	82	PEKLPNYETKIKSFEEHLHTDEEIRYCLE-GSGYFDVRDQNDQWIRALKKG--GMIVL 138	
Db	134	TERGALOELARSEFFWQVL---EAVRHCHNCVGLHRDKDEN---ILIDLNRGELKLIDF 187	
QY	139	PAGMYHRTLTDTDNVYIKAMRLFVGDVPWTPYNRPH 173	
Db	188	GSGALLKDTVTYD--FDGTRVY-SPPEWIRYHRYH 219	
<p>RESULT 8</p> <p>US-08-461-379A-26</p> <p>Sequence 26, Application US/08461379A</p> <p>Patent No. 5871961</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Smith, Kendall A., & Beadling, Carol</p> <p>TITLE OF INVENTION: Nucleic Acids Encoding CR5 Polypeptide,</p> <p>TITLE OF INVENTION: Vector and Transformed Cell Thereof, and</p> <p>TITLE OF INVENTION: Expression Thereof</p> <p>NUMBER OF SEQUENCES: 35</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: Ratner & Prestia (B) STREET:One Westlakes-Berwyn</p> <p>CITY: Valley Forge</p> <p>STATE: Pennsylvania</p> <p>COUNTRY: USA</p> <p>ZIP: 19482</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: PatentIn Release #1.0, Version #1.25</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/461,379A</p> <p>FILING DATE: 5-JUNE-1995</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER: USSN 08/330,108; 08/104,736</p> <p>APPLICATION NUMBER: & 07/796,066</p> <p>FILING DATE: 27-OCT-1994; 10-AUG-1993 & 20-NOV-91</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Viviana Amzel, Ph. D.</p> <p>REGISTRATION NUMBER: 30,930</p> <p>REFERENCE/DOCKET NUMBER: DART-070</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: (610)470-0700</p> <p>TELEFAX: (610)470-0701</p> <p>INFORMATION FOR SEQ ID NO: 26:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 313 amino acids</p> <p>TYPE: peptide</p> <p>STRANDEDNESS: n.a.</p> <p>TOPOLOGY: n.a.</p> <p>MOLECULE TYPE: peptide</p> <p>US-08-461-379A-26</p>			
<p>Y Match 7.4%; Score 81; DB 2; Length 313;</p> <p>Local Similarity 25.2%; Pred. No. 0.5;</p> <p>Matches 39; Conservative 29; Mismatches 55; Indels 32; Gaps</p>			
QY	31	REPKEFIPVDKLT-----LGVISWRLLNPDNW---ENCENLKRREARGSYVDICDVC 81	
Db	85	RVPMEVLLKKVSSGSGVIRLLDWFPERDPSFVLLIERPEVQ	-----DLDFDI 133
QY	82	PEKLPNYETKIKSFEEHLHTDEEIRYCLE-GSGYFDVRDQNDQWIRALKKG--GMIVL 138	
Db	134	TERGALOELARSEFFWQVL---EAVRHCHNCVGLHRDKDEN---ILIDLNRGELKLIDF 187	
QY	139	PAGMYHRTLTDTDNVYIKAMRLFVGDVPWTPYNRPH 173	
Db	188	GSGALLKDTVTYD--FDGTRVY-SPPEWIRYHRYH 219	

3-074B-26
Se 26, Application US/08463074B
No. 6020155

AL INFORMATION:
LICANT: Smith, Kendall A. & Beadling, Carol
LE OF INVENTION: Nucleic Acids Encoding CRI Fusion Protein, Vector an
IBER OF SEQUENCES: 35
'RESPONSE ADDRESS:

ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
ITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

MPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0,
SOFTWARE: Version #1.25
JRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,074B
FILING DATE: 5-JUN-1995

RIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/104,736
FILING DATE: 10-AUG-1993

'RIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/796,066
FILING DATE: 20-NOV-91

VTTORNEY/AGENT INFORMATION:
NAME: Viviana Anzel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38143 (DART-020)

TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 622-7700
TELEFAX: (213) 489-4210

IFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 313 amino acids
TYPE: peptide
STRANDEDNESS: n.a.
TOPOLOGY: n.a.

MOLECULE TYPE: peptide
8-463-074B-26

Very Match 7.4%; Score 81; DB 3; Length 313;
st Local Similarity 25.2%; Pred.No. 0.5;
atches 39; Conservative 29; Mismatches 55; Indels 32; Gaps 9;

31 REPKFPYDKLTE-----LGVISWRNLPNW---ENCENLKIRIAREAGSYVDICDCV 81
| | | : : : | : | : | : | : | :
85 RVPMEVVLLKKVSFGSVIRLLDWFRPDSFVLIERPEPVQ-----DLDFDI 133
| | | : : : | : | : | : | : | :
82 PEKLPNYETKIKSFEEHLHTDEIRICLE-GSGYFDVRCDQMIRIALKKG--GMIVL 138
| | : : | : | : | : | : | : | : | : | :
134 TEGALQEELARLSFFQWLV---EAVRHCHNCGLVHRIDKEN---ILIDLARGEKLIDF 187

139 PAGMYHRTLDTDNYIKAMRLFLVGDPVYPTYNRP 173
| : | : | :
188 GSGALLKDVTYTD--FDGTRVY-SPPWEIRYHYRH 219
| : | : | : | :

JUL 11
U8-468-465-585C-26
sequence 26, Application US/08465585C
Patent No. 6027914

GENERAL INFORMATION:
APPLICANT: Smith, K. A., & Beadling, C.
TITLE OF INVENTION: Nucleic Acids Encoding CR6 Polypeptide, Vecto
TITLE OF INVENTION: Transformed Cell Thereof, and Expression Thereof

SULT 17

SOL-11
-08-465-585C-26
Sequence 26, Application US/08465585C
Patent No. 6027914
GENERAL INFORMATION:
APPLICANT: Smith, K. A., & Beadli
TITLE OF INVENTION: Nucleic Acids
TITLE OF INVENTION: Transformed C

LINE	NUMB
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

,
CORRESPONDENCE ADDRESS:
ADDRESS: PRETTY, SCHROEDER & POPLAWSKI
ADDRESSER: (B) STREET:
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

```

1 ; NUMBER OF SEQUENCES: 35
2 ; CORRESPONDENCE ADDRESS:
3 ; ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
4 ; ADDRESSEE: (B) STREET:
5 ; CITY: Los Angeles
6 ; STATE: California
7 ; COUNTRY: USA
8 ; ZIP: 900071
9 ;
10 ; COMPUTER READABLE FORM:
11 ; MEDIUM TYPE: Floppy disk
12 ; COMPUTER: IBM PC compatible
13 ; OPERATING SYSTEM: PC-DOS/MS-DOS
14 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
15 ; CURRENT APPLICATION DATA:
16 ; FILING APPLICATION NUMBER: US/08/465,585C
17 ; FILING DATE: 5-JUNE-1995
18 ; PRIORITY APPLICATION DATA:
19 ; APPLICATION NUMBER: USSN 08/330,108
20 ; FILING DATE: 27-OCT-1994
21 ; APPLICATION NUMBER: USSN 08/104,736
22 ; FILING DATE: 10-AUG-1993
23 ; APPLICATION NUMBER: USSN 07/796,066
24 ; FILING DATE: 20-NOV-1991
25 ; ATTORNEY/AGENT INFORMATION:
26 ; NAME: Viviana Amzel, Ph. D.
27 ; REGISTRATION NUMBER: 30,930
28 ; REFERENCE/DOCKET NUMBER: P66 38149 (DART-050)
29 ; TELECOMMUNICATION INFORMATION:
30 ; TELEPHONE: (213) 622-7700
31 ; TELEFAX: (213) 4894210
32 ; INFORMATION FOR SEQ ID NO: 26:
33 ; SEQUENCE CHARACTERISTICS:
34 ; LENGTH: 313 amino acids
35 ; TYPE: peptide
36 ; STRANDEDNESS: n.a.
37 ; TOPOLOGY: n.a.
38 ; MOLECULE TYPE: peptide
39 ;
40 ; US-08-465-585C-26
41
42 Query Match 7.4%; Score 81; DB 3; Length 313;
43 Best Local Similarity 25.2%; Pred. No. 0.5;
44 Matches 39; Conservative 29; Mismatches 55; Indels 32; Gaps 9;
45
46 QY 31 REPKEIFDVRLTE-----LGVISWRLNPDNW---ENCENLKRIREARGYSYVDICVC 81
47 DB 85 RYPMVEVLLKKVSSGSGVIRLLDNWFERDPSVLLIERPEVQ-----DLDFDI 133
48 QY 82 PKPLDNVETTKIKSFEBHLTHDEINCYLC- GSGYFDVRDNDOWIRIALKKG--GMIVL 138
49 DB 134 TGRGALQELARSFQVVL---EAVRHCHNCVLVLRDKDEN---ILIDNLRGELKLDIF 187
50 QY 139 PAGMYHREFTLTDNYIKAMRLFVGDVPVTPYNNRH 173
51 DB 188 GSGALKKQTVYTD--FDGTRVY-SPPEWIRYHYH 219
52
53 RESULT 12
54 US-08-652-446-26
55 ; Sequence 26, Application US/08652446
56 ; Patent No. 6057427
57 ; GENERAL INFORMATION:
58 ; APPLICANT: Smith, Kendall A. & Beadling, Carol
59 ; TITLE OF INVENTION: Nucleic Acids Encoding CR5
60 ; TITLE OF INVENTION: Polypeptide, Vector and Transformed Cell Thereof, and
61 ; TITLE OF INVENTION: Expression Thereof
62 ; NUMBER OF SEQUENCES: 38
63 ; CORRESPONDENCE ADDRESS:
64 ; ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
65 ; ADDRESS: (B) STREET:
66 ; CITY: Los Angeles
67 ; STATE: California
68 ; COUNTRY: USA
69 ; ZIP: 900071

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,446
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP App. # 96921319.8
FILING DATE: 5-JAN-1998
APPLICATION NUMBER: PCT/US/96/09194
FILING DATE: 5-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/330,108
FILING DATE: 27-OCT-1994
APPLICATION DATA:
FILING DATE: 5-JUN-1995
APPLICATION NUMBER: 08/463,074
FILING DATE: 5-JUN-1995
APPLICATION DATA:
FILING DATE: 5-JUN-1995
APPLICATION NUMBER: 08/462,337
FILING DATE: 5-JUN-1995
APPLICATION DATA:
FILING DATE: 5-JUN-1995
APPLICATION NUMBER: 08/462,390
FILING DATE: 5-JUN-1995
APPLICATION DATA:
FILING DATE: 5-JUN-1995
APPLICATION NUMBER: 08/465,595
FILING DATE: 5-JUN-1995
APPLICATION DATA:
FILING DATE: 5-JUN-1995
APPLICATION NUMBER: 08/463,081
FILING DATE: 5-JUN-1995
APPLICATION DATA:
FILING DATE: 5-JUN-1995
APPLICATION NUMBER: 08/461,379
FILING DATE: 5-JUN-1991
APPLICATION DATA:
APPLICATION NUMBER: 08/739,523
FILING DATE: 29-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Amzel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: FP66 40035
TELEPHONE: (213) 622-7700
TELEFAX: (213) 489-4210
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 313 amino acids
TYPE: peptide
STRANDEDNESS: n.a.
TOPOLOGY: n.a.
MOLECULE TYPE: peptide
-652-446-26

ry Match 7.4%; Score 81; DB 3; Length 313;
t Local Similarity 25.2%; Pred. No. 0.5;
Matches 39; Conservative 29; Mismatches 55; Indels 32; Gaps 9;
31 REPKEFIPVDKLT-----LGVISWRINPDNW---ENCENLKRIREARGYSYVDICDVC 81
85 RVPMEVLLKKVSSGFSGVIRLLDWFERSFVILIERPEPQV-----DLDFDI 133
82 PEKLPNYETKISFEEHLHTDEIRYCLE-GSGYFDVDRDQNDQWIRALKKG--GMIVL 138
134 TERGAQEELARSFFQVVL---EAVRHCHNCVGLHRDKDEN---ILIDLNRGELKLIDF 187
39 PAGMYHRTFLDTDNYIKAMRLFVGDVPVTPYNRPH 173
88 GSGALLKDTVYTD--FDGTRVY-SPPEWIRYHRYH 219

13
37-543-8

; Sequence 8, Application US/09237543A
; Patent No. 6143540
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosana
; TITLE OF INVENTION: NOVEL MOLECULES OF THE HKID-1-RELATED PROTEIN FAMILY
; FILE REFERENCE: 035800/175631
; CURRENT APPLICATION NUMBER: US/09/237,543A
; CURRENT FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-237-543-8
Query Match 7.4%; Score 81; DB 4; Length 313;
Best Local Similarity 25.2%; Pred. No. 0.5;
Matches 39; Conservative 29; Mismatches 55; Indels 32; Gaps 9;
QY 31 REPKEFIPVDKLT-----LGVISWRINPDNW---ENCENLKRIREARGYSYVDICDVC 81
DB 85 RVPMEVLLKKVSSGFSGVIRLLDWFERSFVILIERPEPQV-----DLDFDI 133
QY 82 PEKLPNYETKISFEEHLHTDEIRYCLE-GSGYFDVDRDQNDQWIRALKKG--GMIVL 138
DB 134 TERGAQEELARSFFQVVL---EAVRHCHNCVGLHRDKDEN---ILIDLNRGELKLIDF 187
QY 139 PAGMYHRTFLDTDNYIKAMRLFVGDVPVTPYNRPH 173
DB 188 GSGALLKDTVYTD--FDGTRVY-SPPEWIRYHRYH 219
RESULT 14
US-09-237-543-9
; Sequence 9, Application US/09237543A
; Patent No. 6143540
; GENERAL INFORMATION:
; APPLICANT: Kapeller, Rosana
; TITLE OF INVENTION: NOVEL MOLECULES OF THE HKID-1-RELATED PROTEIN FAMILY
; FILE REFERENCE: 035800/175631
; CURRENT APPLICATION NUMBER: US/09/237,543A
; CURRENT FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-237-543-9
Query Match 7.4%; Score 81; DB 4; Length 313;
Best Local Similarity 25.2%; Pred. No. 0.5;
Matches 39; Conservative 29; Mismatches 55; Indels 32; Gaps 9;
QY 31 REPKEFIPVDKLT-----LGVISWRINPDNW---ENCENLKRIREARGYSYVDICDVC 81
DB 85 RVPMEVLLKKVSSGFSGVIRLLDWFERSFVILIERPEPQV-----DLDFDI 133
QY 82 PEKLPNYETKISFEEHLHTDEIRYCLE-GSGYFDVDRDQNDQWIRALKKG--GMIVL 138
DB 134 TERGAQEELARSFFQVVL---EAVRHCHNCVGLHRDKDEN---ILIDLNRGELKLIDF 187
QY 139 PAGMYHRTFLDTDNYIKAMRLFVGDVPVTPYNRPH 173
DB 188 GSGALLKDTVYTD--FDGTRVY-SPPEWIRYHRYH 219
RESULT 15
US-09-644-450-8
; Sequence 8

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

in - protein search, using sw model

November 25, 2002, 10:10:00 : Search time 10 seconds
(without alignments)
311.664 Million cell updates/sec

US-09-785-738A-2

score: 1097
e: 1 MENEFOGKTEVIEAWYMD...KEFLAKLLKSEGENQAVEGF 199

table: BLOSUM62
Gapop 10.0 , Gapext 0.5

sd: 100480 seqs, 15651496 residues

number of hits satisfying chosen parameters: 100480

m DB seq length: 0
m DB seq length: 2000000000

rocessing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

18A: Published Applications, AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Hit No.	Score	Query Match	Length	DB ID	Description
1	1097	100.0	199	10	US-09-785-738A-2
2	955.5	87.1	198	10	US-09-785-738A-4
3	862	78.6	200	10	US-09-785-738A-12
4	812	74.0	196	10	US-09-785-738A-6
5	779	71.0	200	10	US-09-785-738A-10
6	691	63.0	189	10	US-09-785-738A-8
7	609	55.3	179	10	US-09-785-738A-16
8	586	53.4	179	10	US-09-840-787-28
9	586	53.4	206	10	US-09-925-301-1387
10	581.5	53.0	181	10	US-09-785-738A-18
11	539.5	49.2	187	10	US-09-785-738A-14
12	129	11.8	181	10	US-09-815-242-5091
13	121.5	11.1	180	10	US-09-815-242-11693
14	95	8.7	153	10	US-09-867-550-458
15	84	7.7	323	10	US-09-971-791-6
16	81	7.4	313	10	US-09-971-791-8
17	81	7.4	313	10	US-09-971-791-9
18	78	7.1	287	10	US-09-912-787-4
19	78	7.1	287	10	US-09-912-787-57

20	78	7.1	287	12	US-10-014-326-43	Sequence 43, Appl
21	78	7.1	287	12	US-10-014-326-68	Sequence 68, Appl
22	78	7.1	305	10	US-09-912-787-6	Sequence 6, Appl
23	78	7.1	305	12	US-10-014-326-70	Sequence 70, Appl
24	78	7.1	497	9	US-09-981-353-105	Sequence 105, App
25	77	7.0	313	10	US-09-971-791-7	Sequence 7, Appl
26	74.5	6.8	540	12	US-10-022-461-2	Sequence 2, Appl
27	74	6.7	803	9	US-09-968-436B-2	Sequence 2, Appl
28	74	6.7	803	10	US-09-759-010-7	Sequence 7, Appl
29	70.5	6.4	302	10	US-09-815-242-5495	Sequence 5495, Ap
30	70.5	6.4	312	10	US-09-815-242-12432	Sequence 12432, A
31	70	6.4	316	10	US-09-462-846-5	Sequence 5, Appl
32	70	6.4	1534	9	US-09-736-968A-10	Sequence 10, Appl
33	70	6.4	1534	10	US-09-736-969A-10	Sequence 10, Appl
34	70	6.4	1534	10	US-09-736-960-10	Sequence 70, Appl
35	69.5	6.3	4861	10	US-09-919-497-70	Sequence 10, Appl
36	69	6.3	696	10	US-09-815-242-5443	Sequence 5443, Ap
37	69	6.3	698	10	US-09-815-242-12325	Sequence 12325, A
38	68.5	6.2	544	10	US-09-887-389-2	Sequence 2, Appl
39	68	6.2	539	10	US-09-965-313-6	Sequence 6, Appl
40	67.5	6.2	668	10	US-09-925-300-1677	Sequence 1677, Ap
41	67	6.1	357	10	US-09-907-537-2	Sequence 2, Appl
42	67	6.1	498	10	US-09-815-242-10556	Sequence 10556, A
43	67	6.1	500	10	US-09-815-242-5007	Sequence 5007, Ap
44	66.5	6.1	145	10	US-09-864-761-44376	Sequence 44376, A
45	66.5	6.1	200	10	US-09-966-608-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-785-738A-2
; Sequence 2, Application US/09785738A
; Patent No. US20020032918A1
; GENERAL INFORMATION:
; APPLICANT: Sauter, Margaret M
; APPLICANT: Lorblecke, Rene
; TITLE OF INVENTION: ALTERATION OF GROWTH AND ADAPTATION UNDER HYPOXIC
; FILE OF INVENTION: CONDITIONS
; FILE REFERENCE: 2283/201
; CURRENT APPLICATION NUMBER: US/09/785,738A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,572
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Rice
US-09-785-738A-2

Query Match	100.0%	Score	1097	DB	10	Length	199
Best Local Similarity	100.0%	Pred. No.	4.2e-102	Mismatches	0	Indels	0
Matches	199	Conservative	0	Matches	0	Gaps	0
QY	1	MENEFOGKTEVIEAWYMDSEEDQRLPHHREPKEFPVDKLTCLGVISWRLNPDNWC	60				
DB	1	MENEFOGKTEVIEAWYMDSEEDQRLPHHREPKEFPVDKLTCLGVISWRLNPDNWC	60				
QY	61	ENLKRIRARGSYVDCVCPKLPNYETKIKSPFEEHLHTDEIRYCLEGSGYFVD	120				
DB	61	ENLKRIRARGSYVDCVCPKLPNYETKIKSPFEEHLHTDEIRYCLEGSGYFVD	120				
QY	121	QNDQWIRIALKKGMIVLPAAGMYHRTFLDNDYIKAMRLFVGDVPMYTPYNRPHDL	180				
DB	121	QNDQWIRIALKKGMIVLPAAGMYHRTFLDNDYIKAMRLFVGDVPMYTPYNRPHDL	180				
QY	181	EFLAKLLKSEGENQAVEGF	199				
DB	181	EFLAKLLKSEGENQAVEGF	199				

1

```
TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 179 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LM020703
; CLONE: 1577239
; SEQUENCE DESCRIPTION: SEQ ID NO: 28 :
US-09-840-787-28

Query Match 53.4%; Score 586; DB 10; Length 179;
Best Local Similarity 56.7%; Pred. No. 2.7e-51;
Matches 101; Conservative 34; Mismatches 43; Indels 0; Gaps 0;

y 12 VIEAWYMDSEEDQRLPHHREPKEFIPVDKLTGVLISWRLNPDNWCENLKRREARG 71
b 1 MVQAWYMDADPGDQRPQHPDPGRPVGLEQLRLGLVYKLDADKYENDPELEKIRERN 60
y 72 YSYVDICDVCPEKLPNYETKISFEEHLHTDEIRYCLGSGYFDVRDQNDQWIRIALK 131
> 61 YSWMDIITICKDKLPNYEEKIKMFEYEEHLHLDDEIRYILDGSGYFDVRDQNDQWIRIFME 120
, 132 KGMIVLPAGMYHRETTDTDNYYIKAMRLFVGPVWTPYRNPHDHLPAKKEFLAKLKS 189
121 KGMVTLPAIGYHRETTVDENYTKAMRLFVGPVWTPYRNPAHDFEARGQYVVKFLAQT 178

SULT 9
-09-925-301-1387
Sequence 1387; Application US/09925301
Patent No. US20020032918A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05882
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1694
SOFTWARE: PatentIn Ver. 2.0
Q ID NO 1387
LENGTH: 206
TYPE: PRT
ORGANISM: Homo sapiens
US-925-301-1387

ry Match 53.4%; Score 586; DB 10; Length 206;
t Local Similarity 56.7%; Pred. No. 3.3e-51;
ches 101; Conservative 34; Mismatches 43; Indels 0; Gaps 0;

12 VIEAWYMDSEEDQRLPHHREPKEFIPVDKLTGVLISWRLNPDNWCENLKRREARG 71
28 MVQAWYMDADPGDQRPQHPDPGRPVGLEQLRLGLVYKLDADKYENDPELEKIRERN 87
72 YSYVDICDVCPEKLPNYETKISFEEHLHTDEIRYCLGSGYFDVRDQNDQWIRIALK 131
88 YSWMDIITICKDKLPNYEEKIKMFEYEEHLHLDDEIRYILDGSGYFDVRDQNDQWIRIFME 147
132 KGMIVLPAGMYHRETTDTDNYYIKAMRLFVGPVWTPYRNPHDHLPAKKEFLAKLKS 189
148 KGMVTLPAIGYHRETTVDENYTKAMRLFVGPVWTPYRNPAHDFEARGQYVVKFLAQT 205

10
85-738A-18
nce 18; Application US/09785738A
t No. US20020032918A1
```

```
GENERAL INFORMATION:
; APPLICANT: Sauter, Margret M
; APPLICANT: Lorbiecke, Rene
; TITLE OF INVENTION: ALTERATION OF GROWTH AND ADAPTATION UNDER HYPOXIC
; TITLE OF INVENTION: CONDITIONS
; FILE REFERENCE: 2283/201
; CURRENT APPLICATION NUMBER: US/09/785,738A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,572
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Zebrafish
; OTHER INFORMATION: n at positions 634 and 642 is unknown
US-09-785-738A-18

Query Match 53.0%; Score 581.5; DB 10; Length 181;
Best Local Similarity 59.1%; Pred. No. 7.8e-51;
Matches 104; Conservative 28; Mismatches 43; Indels 1; Gaps 1;

QY 12 VIEAWYMD-0SEEDQRLPHHREPKEFIPVDKLTGVLISWRLNPDNWCENLKRREARG 70
Db 3 VFEAWYMDSEEDQRLPHKLSNPQVSVQQLGHEIGVPHKLNADYENDPELEKIREK 62
QY 71 GYSYVDICDVCPEKLPNYETKISFEEHLHTDEIRYCLGSGYFDVRDQNDQWIRIAL 130
Db 63 GYSFMDIITIHDPKLPDYQNKLMFYEEHLHLDDEIRYILEGSSYFDVRDEGDRWIRIAV 122
QY 131 KKGSMIVLPAGMYHRETTDTDNYYIKAMRLFVGPVWTPYRNPHDHLPAKKEFLAKL 186
Db 123 SKGDLITLPAGIYHRETTVDENYTKAMRLFVGPVWTPYRNPAHDFEARGQYVVKFLAQT 178

RESULT 11
US-09-785-738A-14
; Sequence 14; Application US/09785738A
; Patent No. US20020032918A1
GENERAL INFORMATION:
APPLICANT: Sauter, Margret M
APPLICANT: Lorbiecke, Rene
TITLE OF INVENTION: ALTERATION OF GROWTH AND ADAPTATION UNDER HYPOXIC
TITLE OF INVENTION: CONDITIONS
FILE REFERENCE: 2283/201
CURRENT APPLICATION NUMBER: US/09/785,738A
CURRENT FILING DATE: 2001-02-16
PRIOR APPLICATION NUMBER: 60/183,572
PRIOR FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 187
TYPE: PRT
ORGANISM: Human
US-09-785-738A-14

Query Match 49.2%; Score 539.5; DB 10; Length 187;
Best Local Similarity 55.1%; Pred. No. 1.2e-46;
Matches 98; Conservative 29; Mismatches 50; Indels 1; Gaps 1;

QY 12 VIEAWYMDSEEDQRLPHHREPKEFIPVDKLTGVLISWRLNPDNWCENLKRREARG 71
Db 10 VQAWYMDADPGDQRPQHPDPGRPVGLEQLRLGLVYKLDADKYENDPELEKIRERN 68
QY 72 YSYVDICDVCPEKLPNYETKISFEEHLHTDEIRYCLGSGYFDVRDQNDQWIRIALK 131
Db 69 YSWMDIITICKDKLPNYEEKIKMFEYEEHLHLDDEIRYILDGSGYFDVRDQNDQWIRIFME 128
QY 132 KGMIVLPAGMYHRETTDTDNYYIKAMRLFVGPVWTPYRNPHDHLPAKKEFLAKL 186
Db 129 KGMVTLPAIGYHRETTVDENYTKAMRLFVGPVWTPYRNPAHDFEARGQYVVKFLAQT 178
```

Query Match 8.7%; Score 95; DB 10; Length 153;

Best Local Similarity 25.9%; Pred. No. 0.017;
Matches 37; Conservative 20; Mismatches 64; Indels 22; Gaps 5;

28 PHHREPKEFIPVDKLTGLVISWR-----LNPDNWCENLKRREA-----RGYS 73

14 PELRNDPDVIASRLREIGVRFERWASPVALSPD-----DPAETILEAYRPHLDALMGES 68

74 YVDICDV--CPEKLPNYETKIKSFFEEHLHTDEEIRCYCLESGYFDVRDQNDQWIRALK 131 QY

69 GAGSADVIKISAATPNKOALRDKFLSEHOHTEDEVRFVHGSGNF-IMHVDGRVYDAHCT 127

Qy 132 KGGMIVLPAGMYHRTLDTDNYI 154

```

:| :| :|| | | | :|
128 OGDLSVPASTKHWFDAGPDPFV 150
00b

```

RESULT 15

IS-09-971-791-6

Sequence 6, Application US/09971791

Patent No. US20020115120A1

GENERAL INFORMATION:

APPLICANT: Rosanna Kapeller-Libermann

APPLICANT: Laura A. Rudolph-Owen

APPLICANT: Kyle MacBeth
TITLE OF INVENTION: NOVEL MOLECULES OF THE HKID-1-RELATED PROTEIN FAMILY AND USES THEREFOR

FILE REFERENCE: 35800/238856

CURRENT APPLICATION NUMBER: US/09/971,791

CURRENT FILING DATE: 2001-10-04

PRIOR APPLICATION NUMBER: 09/644,450

PRIOR FILING DATE: 2000-08-23

PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: 09/237,543

PRIOR FILING DATE: 1999-01-26

PRIOR FILING DATE: 1933
 NUMBER OF SEQ ID NOS: 11

NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSEQ for Windows Version 4.0

,EQ ID NO 6

SEQ ID NO: 323

LENGTH: 52
TYPE: PRT

TYPE: PRT
ORGANISM: *Xenopus laevis*

ORGANISM: XE
09-971-791-6

Very Match

very Match
7.78; Score 84; DB 10; Length 323;

Very Match 7.78; Score 84; DB 10; Bengon 323;
3st Local Similarity 24.28; Pred. No. 0.55;

33 KFEFIPVDKLTE-----LGVISWRLNPDNW-----ENCENLKRIREARGYSYVDICDVCPE 83

[illegible]

84 KLPNYETKIKSFFEEHLTDEEIRYCLE-GSGYFVDRDQNDQWIRIALKKG--GMIVLPA 140

[illegible]

141 GMYHRETLDTDNYIKAMRLFVGDPVWTPYNRPH 173

h completed: November 25, 2002, 10:13:50

```

11 completed: 1
time : 11 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

protein search, using sw model

November 25, 2002, 10:08:00 ; Search time 20 Seconds
(without alignments)
956,537 Million cell updates/sec

US-09-785-738A-2
1 MENEFDQKTEVTEAMND.....KEFLAKLLKSEGENQAVEGF 199

BLOSUM62
Gapop 10.0 , Gapext 0.5

283224 seqs, 96134422 residues
283224

DB seq length: 0
DB seq length: 2000000000
Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

PIR 73:3
1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Seq.	Score	Query Match %	Length	DB	ID	Description
1	1097	100.0	199	2	T02787	probable submergen
2	955.5	87.1	198	2	T02918	probable submergen
3	789.5	72.0	940	2	H71409	hypothetical prote
4	781	71.2	199	2	T00973	hypothetical prote
5	362	33.0	179	2	S53039	probable nickel-bi
6	307.5	28.0	178	2	T40726	probable nickel-bi
7	307.5	28.0	178	2	T32103	hypothetical prote
8	229	20.9	221	2	T36578	hypothetical prote
9	227	20.7	159	2	T29472	hypothetical prote
10	192	17.5	178	2	B69884	probable methionin
11	148	13.5	188	2	F82587	dioxygenase XF2210
12	138	12.6	183	2	E70469	probable methionin
13	129	11.8	181	2	E83436	probable oxidase P
14	121.5	11.1	182	2	A62146	hypothetical prote
15	118	10.8	180	1	A59159	methionine salvage
16	91	8.3	399	2	B70349	hypothetical prote
17	90	8.2	659	2	A75029	h+-transporting Ar
18	87.5	8.0	307	2	A42902	transcription regu
19	87.5	8.0	307	2	P83198	transcription regu
20	86.5	7.9	264	2	S49787	hypothetical prote
21	85.5	7.8	465	2	H86482	protein F5J5.11 [i
22	85.5	7.8	286	2	A54070	3-hydroxyanthranil
23	85.5	7.8	315	2	D84139	mannose-6-phosphat
24	85.5	7.8	1279	2	E84709	type IIS restricti
25	83.5	7.6	4273	2	C69679	polyketide synthas
26	82	7.5	518	1	FWOAG1	12S seed storage g
27	81.5	7.4	136	2	C97153	probable transcrip
28	81.5	7.4	337	2	F70323	transcription term
29	81	7.4	313	1	TVHUP1	protein kinase (EC

30	81	7.4	313	1	S26298	protein kinase (EC
31	81	7.4	508	2	S35692	UTP-glucose-1-phos
32	80.5	7.3	502	2	S35158	triticin - wheat
33	80.5	7.3	514	2	T39458	DNA binding protei
34	80	7.3	360	1	S41759	ribosomal protein
35	80	7.3	311	2	D97058	SNW-dependent meth
36	80	7.3	493	2	S17762	glutelin glub-1 pr
37	80	7.3	795	2	S33101	H4L protein - vari
38	80	7.3	795	2	T28525	hypothetical prote
39	80	7.3	795	2	E72161	J4L protein - vari
40	79.5	7.2	709	2	S40926	hypothetical prote
41	79.5	7.2	784	2	C88558	protein ZK1098.3 [
42	79.5	7.2	1471	1	S30790	myosin My04 - yeas
43	79	7.2	575	1	QOVZHS	H5 protein - vacci
44	79	7.2	714	2	T22454	hypothetical prote
45	79	7.2	795	1	QOVZHS	H4L protein - vacc

ALIGNMENTS

RESULT 1

T02787
probable submergence induced protein 2 - rice
C:Species: Oryza sativa (rice)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
C:Accession: T02787
R:Lorbiecke, R.; Sauter, M.
submitted to the EMBL Data Library, February 1998
A:Reference number: Z14738
A:Accession: T02787
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-199 <LOR>
A:Cross-references: EMBL:AF050200; NID:g2952337; PID:g2952338
C:Genetics:
A:Gene: SIP2

Query Match 100.0%; Score 1097; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 6e-88;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MENEFDQKTEVTEAMNDSEEDQRLPHRPEKPEIPVDKLTGLGVISMRKLPDNNWNC	60
Db	1	MENEFDQKTEVTEAMNDSEEDQRLPHRPEKPEIPVDKLTGLGVISMRKLPDNNWNC	60
Qy	61	ENLKRIREARGSYVDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLESGYFDVRD	120
Db	61	ENLKRIREARGSYVDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLESGYFDVRD	120
Qy	121	QNDQWIRIALKKGGMIVLPACGMVHRFTLDTDNYIKAMRLFVGDVPTVTPNRPDHLPAK	180
Db	121	QNDQWIRIALKKGGMIVLPACGMVHRFTLDTDNYIKAMRLFVGDVPTVTPNRPDHLPAK	180
Qy	181	EFLAKLLKSEGENQAVEGF	199
Db	181	EFLAKLLKSEGENQAVEGF	199

RESULT 2

T02918
probable submergence induced, nickel-binding protein 2A - rice
C:Species: Oryza sativa (rice)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C:Accession: T02918
R:Lorbiecke, R.; Sauter, M.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z14768
A:Accession: T02918
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-198 <LOR>
A:Cross-references: EMBL:AF068332; NID:g3201968; PID:g3201969


```

Query Match          28.0%; Score 307.5; DB 2; Length 178;
Best Local Similarity 34.4%; Pred. No. 1.8e-19;
Matches 63; Conservative 37; Mismatches 66; Indels 17; Gaps 27;

QY 15 ANYWDSDEQDRLPHRH-EPKEFIPVDKLTGELGVISWRNLNPONWNCENLKRIBEARQYS 73
DB 2 A YFDTVTNKKRQDECRSPNKATEOLLRIGVCTKVFNDQEHVQDGLRLKKYDMN 61

QY 74 YVDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLEGSGYFDVRDQNDQWIRIALKKG 133
DB 62 FRDEVHICRATFMENDEKLDIFEFELHDDAELRVTKHGVGFDVTRKDEAMIRIPVRG 121

QY 134 GNVILPAGMYHRETTLDTONYIKAMLEFGDVPVTPNRPDHPILPARKEFLAKLLKSEGN 193
DB 122 DFEVLPAGIYHRETTDPSDDVVALFRNNPKWTAFNR-----KADGDE 165

QY 194 QAV 196
DB 166 QRV 168

RESULT 8
T16578
hypothetical protein K07E1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16578
R:Fulton, B.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of C. elegans cosmid K07E1.
A:Reference number: T18539
A:Accession: T16578
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-221 <PUL>
A:Cross-references: EMBL:U23173; NID:g726396; PID:g726397; PIDN:AAC46708.1; CES:
A:Experimental source: strain Bristol N2
A:Genetics:
A:Gene: CESP:K07E1.1
A:Introns: 15/3; 206/3

Query Match          20.9%; Score 229; DB 2; Length 221;
Best Local Similarity 33.1%; Pred. No. 1.5e-12;
Matches 53; Conservative 33; Mismatches 70; Indels 4; Gaps 4;

QY 7 DOKTEVTENWYMDSE-EDORLPFHREKEFIPVDKLTE-LGVISWRNLNP-DNWNCENL 63
DB 59 DFQLKMWIOMQMEPYCGDPRLPHLFPFKKITPDELSKRTGLYWKLTLDLQVALAKRL 118

QY 64 KRIBEARQYSYVDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLEGSGYFDVRDQND 123
DB 119 TTKLEHSFKKEDIETLDAETANFDDXIEELFESSVPFEQARNIIEGTATYVDKNG 178

QY 124 QWIRIALKKGMYLPGMYHRETTLDTONYIKAMRLEFGD 163
DB 179 QWVRIFCEYGLIILIPANTCFRFTTTPHNFVK-MRRFYKD 217

RESULT 9
T29472
hypothetical protein T01D1.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29472
R:Bradshaw, H.; Wohlmann, P.
submitted to the EMBL Data Library, November 1996
A:Description: The sequence of C. elegans cosmid T01D1.
A:Reference number: Z20623
A:Accession: T29472
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-159 <BRA>

```

Db 117 FVIOQDCTFDVRLNPGDILISVPENIRHYFTLQEDRKVAVRIVTTGGWP 169

RESULT 11

F82587

dioxygenase XF2210 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000

C:Accession: F82587

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: F82587

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-188 <SIM>

A:Cross-references: GB:AE004033; GB:AE003849; NID:g9107342; PIDN:AAEP5009.1; GSPDB:GN

A:Experimental source: strain 9a5c

R:Simpson A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R

Briones, M.R.S.; Bueno, C.; El-Dorfi, H.; Facincan, A.P.; Ferreira, A.J.S.

as-Neto, E.; Docena, C.; El-Dorfi, H.; Facincan, A.P.; Ferreira, A.J.S.

submitted to Genbank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.D.; Marques, M.V.; Martins

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Mencia, C.F.M.; Miracca, E.C.; Miyaki, C.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv

M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF2210

C:Superfamily: Klebsiella methionine salvage pathway enzyme E-2/E-2'

Query Match 13.5%; Score 148; DB 2; Length 188;

Best Local Similarity 30.5%; Pred. No. 1.3e-05;

Matches 32; Conservative 17; Mismatches 54; Indels 2; Gaps 2;

QY 62 NLKRIEARGSYVYDVCVCEPKLPNYETKIKSPFEHLHTDEIRYCLEGGSYFQVDRDQ 121

Db 61 DIDRLIATHGFKTVVVSIAPON-PKREEMRAKLEEHFKEDVRFVAGSLGTFVHS 119

QY 122 NDOWIRIALKGGMIVLPAGMYHRTFTDPTDNYIKAMFLVCGDPW 166

Db 120 NKVY-ETECYKNDLIAIPDGTQHWFDGMAAPYFAIRFETPDGW 163

RESULT 12

E70469

probable methionine salvage pathway enzyme E-2/E-2' aq_1975 [similarity] - Aquifex ae

N:Contains: 5-methylthio-3-oxo-1-penten-1,2-diol dioxygenase [EC 1.13.11.1.]

C:Species: Aquifex aeolicus

C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 21-Jul-2000

C:Accession: E70469

R:Becker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;

Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: E70469

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-183 <AQF>

A:Cross-references: GB:AE000766; NID:g2984216; PIDN:AAC07759.1; PID:g2984230; GB:AE00

A:Experimental source: strain VF5

C:Genetics:

A:Gene: aq_1975

C:Superfamily: Klebsiella

GenCore version 5.1.3

L

November 25, 2002, 10:00:25 : Search time 12 seconds

(wit

687.816 Million cell

US-09-785-738A-2

score: 1097

table: BIOSUM62

Gapop 10.0 , Gapext 0.5

ad: 112892 seqs, 41476328 residues

[illegible]

number of hits satisfying chosen parameters:

DB seq length: 0

```
DB seq length: 2000000000
DB seq length: 2000000000
```

processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 su

Red and Blue 10: 4

use : SWANSTRUC_401.3

SUMMARIES

ID	DB	Length	Query Match	Score	%		Description
1	1	179	33.0	362	1	YMO9_YEAST	Q03677 saccharomyc
2	1	399	8.3	91	1	Y545_AQUAE	Q06822 aquifex aeo
3	1	659	8.2	90	1	VAT1_PYRAB	Q09ux2 pyrococcus
4	1	307	8.0	87.5	1	MMGR_PSEAE	P28809 pseudomonas
5	1	264	7.9	86.5	1	Y1J3_YEAST	P40496 saccharomyc
6	1	286	7.8	85.5	1	3HAO_HUMAN	P64952 homo sapien
7	1	295	7.7	85	1	RS1_HYMP7	Q03698 synchoboccc
8	1	323	7.7	84	1	PIM3_XENLA	Q91822 xenopus lae
9	1	750	7.6	83.5	1	APB2_RAT	Q035431 rattus norv
10	1	4273	7.6	83.5	1	PKSM_BACSU	P40872 bacillus su
11	1	318	7.6	83	1	PIM1_BOVIN	Q9n0p9 bos taurus
12	1	518	7.5	82	1	SSGI_AVESE	P12615 avena sativ
13	1	313	7.4	81	1	PIM1_FELCA	Q951j0 felis silve
14	1	313	7.4	81	1	PIM1_HUMAN	P11309 homo sapien
15	1	313	7.4	81	1	PIM1_RAT	P26794 rattus norv
16	1	507	7.4	81	1	UDP1_CRIGR	Q035156 cricetulus
17	1	507	7.4	81	1	UDP1_HUMAN	Q07131 homo sapien
18	1	311	7.3	80	1	PRMA_CLOAB	P45558 clostridium
19	1	499	7.3	80	1	GLUB_ORISA	Q02898 oryza sativ
20	1	795	7.3	80	1	RP94_VARV	P33067 variola vir
21	1	784	7.2	79.5	1	YOG3_CAEEL	P34603 caenorhabdi
22	1	1471	7.2	79.5	1	MY34_YEAST	P32492 saccharomyc
23	1	286	7.2	79	1	RP94_VACCV	P07241 vaccinia vi
24	1	295	7.2	78.5	1	3HAO_PSP1	P46953 rattus norv
25	1	237	7.1	78	1	GP44_BPS1	Q08402 bacterioph
26	1	497	7.1	78	1	UDP2_HUMAN	Q16851 homo sapien
27	1	387	7.1	77.5	1	WR36_ARATH	Q9car4 arabidopsis
28	1	749	7.1	77.5	1	APB2_HUMAN	Q95767 homo sapien
29	1	313	7.0	77	1	PIM1_MOUSE	P06803 mus musculu
30	1	323	7.0	77	1	PIW3_COTJA	Q9pu85 coturnix co
31	1	507	7.0	77	1	UDP2_PIG	P79303 sus scrofa
32	1	1755	7.0	77	1	YJ29_YEAST	P40710 saccharomyc
33	1	507	6.9	76	1	UDP2_BOVIN	Q07130 bos taurus

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M84911; AAA25890.1; -
EMBL; AE004778; AAG06959.1; -
PIR; S27601; S27601.
PIR; A42902; A42902.
InterPro; IPR000005; HTHARAC.
Pfam; PF00165; HTH_Arac; 2.
PRINTS; PR00032; HTHARAC.
SMART; SM00342; HTH_ARAC; 1.
PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
Transcription regulation; DNA-binding; Activator; Complete proteome.
DNA_BIND 217 236 H-T-H MOTIF (BY SIMILARITY).
CONFLICT 129 129 S -> T (IN REF. 1).
SEQUENCE 307 AA; 35379 MW; 9BF7AEE348FAFD45 CRC64;

Query Match 8.0%; Score 87.5; DB 1; Length 307;
Best Local Similarity 28.2%; Pred. No. 1.8;
Matches 20; Conservative 10; Mismatches 40; Indels 1; Gaps 1;

79 DVCPKLPNVEYTKSPFEHLHTDEIRYCLGSGVDFVRDQNDQWIRIALKKGGMIVL 138

44 DLPTSGLHYRRARDHMRSEHDEHLIYCSEGGQLLRVR-EGEARVYRSGDLLWL 102

139 PAGMYHRTLD 149

103 PPGMAHYAAD 113

T 5

YEAST
P40496;
01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Hypothetical 30.5 kDa protein in SGAL-KTR7 intergenic region.
Y11093C.
Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
[1]

SEQUENCE FROM N.A.
STRAIN=S288c / AB972;
Barrell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
Louis E., Lye G., Moulle T., Moutle T., Odell C., Pearson D.,
Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
Walsh S.V., Whitehead S.;
Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Z46728; CAA86701.1; -
SGD; S0001355; Y11093C.
Hypothetical protein.
SEQUENCE 264 AA; 30513 MW; D6E5B8B81A2B6C2E CRC64;

Query Match 7.9%; Score 86.5; DB 1; Length 264;
Best Local Similarity 23.2%; Pred. No. 1.9;
Matches 48; Conservative 29; Mismatches 75; Indels 55; Gaps 11;

QY 6 QDGKTEVIAWY-----MDSSEDOPLPHHREPKEEIPVDKLTTEL-----GVISWR 51

DB 19 KSGLLKETFAWNVVASIPSTKFTREPRKPNPSNGHIIKGLVDVTEQPHANNKGLYKTR 78

QY 52 LNPDNWENCENLKRIAREAGYSYVDICDVCPEKLPNYETKIKS-FFEEHLHTDEEIRYCL 110

DB 79 PNSSD-----KRGVGRLYR-----PPKLTVEDRLSLFVKQHPWELSRPKILV 123

QY 111 EGSYGFVDRDQNDQW---IRIALKKGGMIVLPAGNYHRTLDTDNYIKAMRLFVGDVWVT 167

DB 124 EN-----EIGDENYDWSHMLQIGRPLDGESVIQRTMY---LIKTKQY-----GDMV-- 166

QY 168 PYNRPDHLPARKEFLAKLLKSEGENQ 194

DB 167 ---EAYDH---ARVEFYALRMOEETEQQ 188

RESULT 6

3HAO_HUMAN
ID 3HAO_HUMAN STANDARD; PRT; 286 AA.
AC P46952;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-hydroxyanthranilate 3,4-dioxygenase (EC 1.13.11.6) (3-HAO)
DE (3-hydroxyanthranilic acid dioxygenase) (3-hydroxyanthranilate
dioxygenase).
GN HAAO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=94245687; PubMed=7514594;
RA Malherbe P., Kohler C., da Prada M., Lang G., Kiefer V.,
RA Schwarcz R., Lahm H., Cesura A.M.;
RT "Molecular cloning and functional expression of human 3-
RT hydroxyanthranilic-acid dioxygenase";
RL J. Biol. Chem. 269:13792-13797(1994).
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF THE EXCITOTOXIN QUINOLINIC
CC ACID (QUIN) FROM 3-HYDROXYANTHRANILIC ACID. THE DIRECT PRODUCT
CC OF THE REACTION SPONTANEOUSLY REARRANGE TO QUIN.
CC -!- CATALYTIC ACTIVITY: 3-hydroxyanthranilate + O(2) -> 2-amino-3-
CC carboxymuconate semialdehyde.
CC -!- COFACTOR: FERROUS ION.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; Z29481; CAA82618.1; -
DR Genew; HGNC:4796; HAAO.
DR MIM; 604521; -
KW Oxidoreductase; Dioxygenase; Iron.
SQ SEQUENCE 286 AA; 32542 MW; 4DA10F20FC635885 CRC64;

Query Match 7.8%; Score 85.5; DB 1; Length 286;
Best Local Similarity 24.1%; Pred. No. 2.5;
Matches 34; Conservative 25; Mismatches 37; Indels 45; Gaps 7;

QY 45 LGVISWRLPDPDNWENCENLKRIAREAGYSYVDICDVCPEKLPNYETKIKSFF----- 96

[illegible][illegible]

```
Subtilist: BG10931; pksm.
InterPro: IPR002198; ADH_short.
InterPro: IPR000794; ketoacyl-synt.
InterPro: IPR001601; Methyltransf.
InterPro: IPR003880; Pantne_attach.
InterPro: IPR000051; SAM_bind.
Pfam: PF00106; adh_short; 1.
Pfam: PF00109; ketoacyl-synt; 3.
Pfam: PF00550; pp-binding; 4.
Pfam: PF02801; ketoacyl-synt_C; 3.
PROSITE: PS00012; PHOSPHOPANTHETHEINE; 2.
PROSITE: PS00606; B_KETOACYL_SYNTHASE; 2.
PROSITE: PS50075; ACP_DOMAIN; 4.
Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP;
Phosphopantetheine; Multifunctional enzyme; Repeat; Complete proteome.
DOMAIN 295 364
DOMAIN 396 834
DOMAIN 2190 2258
DOMAIN 2322 2737
DOMAIN 3532 3947
DOMAIN 3410 3483
DOMAIN 4140 4209
DOMAIN 327 327
BINDING 2222 2222
ACT_SITE 2476 2476
ACT_SITE 3446 3446
ACT_SITE 3690 3690
ACT_SITE 4172 4172
CONFLICT 103 103
CONFLICT 276 276
CONFLICT 289 289
SEQUENCE 4273 AA; 477459 MW; 3BBFCF1A250AEB5A CRC64;

Query Match 7.6%; Score 83.5; DB 1; Length 4273;
Best Local Similarity 22.2%; Pred. No. 74;
Matches 54; Conservative 24; Mismatches 64; Indels 101; Gaps 11;

11 EVTEAWMDSEEDQLPH-----HREP-----KEFIPVDKL 42
12 ELIEAYRDDASQAQPEAYVPALEPEASKSIRQVHDEPIAIGSGRPOQDSV 3547
43 TELGVISWRLNPDWNCENLKRI-----REAGYSY-----VDICD 79
3548 HEL-----WDNLKNGKSCISDIPGERDWDGKRNDRPEKAVPRMGAFKLKIDRFED 3596
80 -----VCPEKLPNYETKIKSFFEEHLHTDEIRYCLEGSGYFVDRONDQWIRALKKG 134
3597 PLFFQISPKAESMDPRQPIFLEAWHTFEDAGMGD-----RIKGSKG 3641
135 MIV-LPAGMYHRTLDTDNYIKAMRLFVGDVPVTPYNRPHDLPARKEFLAKLKSGEN 193
3642 VYGVVEGEYAHLTGDTD-YINGTQ-----NATLSAR---IAYALDLKGP 3683
194 QAV 196
3684 MAL 3686

RESULT 11
IML_BOVIN STANDARD; PRT; 313 AA.
D PML_BOVIN
C QN0P9;
T 15-JUN-2002 (Rel. 41, Created)
T 15-JUN-2002 (Rel. 41, Last sequence update)
T 15-JUN-2002 (Rel. 41, Last annotation update)
E Proto-oncogene serine/threonine-protein kinase pim-1 (EC 2.7.1.37).
N PIM1.
S Bos taurus (Bovine).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
C Bovidae; Bovinae; Bos.
X NCBI_TaxID=9913;
N [1]
```

```
SEQUENCE FROM N.A.
RX MEDLINE=21109090; PubMed=11182156;
Wang Z., Petersen K., Weaver M.S., Magnuson N.S.;
"CDNA cloning, sequencing and characterization of bovine pim-1.";
Vet. Immunol. Immunopathol. 78:177-195(2001).
RL CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -|- PTM: Autophosphorylated (By similarity).
CC -|- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC PIM SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: AF259078; AAF67200.1; -.
HSP: Q63450; 1A06.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Proto-oncogene;
KW ATP-binding; phosphorylation.
FT DOMAIN 38 290 PROTEIN KINASE.
FT NP_BIND 44 52 ATP (BY SIMILARITY).
FT BINDING 67 67 ATP (BY SIMILARITY).
FT ACT_SITE 167 167 BY SIMILARITY.
FT ACT_SITE 167 167 BY SIMILARITY.
SQ SEQUENCE 313 AA; 35629 MW; 9EF40229A847AD47 CRC64;

Query Match 7.6%; Score 83; DB 1; Length 313;
Best Local Similarity 25.2%; Pred. No. 4.6;
Matches 39; Conservative 30; Mismatches 54; Indels 32; Gaps 9;

31 REKEFTVPDKLTE-----LGVISWRLNPDW-----NCENLKRIREARGYSYVDICDVC 81
85 RVPWEVLLKKVSGSGVIRLLDWFEPDSFVLLERPEPVQ-----DLDFDI 133
82 PEKLPNYETKIKSFFEEHLHTDEIRYCLE-GSGYFVDRONDQWIRALKKG-GMIVL 138
134 TERGAQEELARSEFFQVQL--EAVRHCHDGVLRHDKDEN---ILIDLNRGELKLIDF 187
139 PAGMYHRTLDTDNYIKAMRLFVGDVPVTPYNRPH 173
188 GSGALLKDTVTVD--FDGTRVY-SPPEWIRYRHYH 219

RESULT 12
SSGI_AVEA STANDARD; PRT; 518 AA.
AC P12615;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE 01-FEB-1994 (Rel. 28, Last annotation update)
DE 12S seed storage globulin precursor.
OS Avena sativa (Oat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
OC Avenaceae; Avena.
OX NCBI_TaxID=4498;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams;
RA Shotwell M.A., Afonso C., Davies E., Chesnut R.S., Larkins B.A.;
RT "Molecular characterization of oat seed globulins.";
RL Plant Physiol. 87:698-704(1988).
CC -|- FUNCTION: THIS IS A SEED STORAGE PROTEIN.
CC -|- SUBUNIT: HEXAMER; EACH SUBUNIT IS COMPOSED OF AN ACIDIC AND A
```

***Comparison of the human and mouse PIM-1 cDNAs: nucleotide sequence**

protein.";
Oncogene Res. 1:103-112(1987).
[4]
SEQUENCE FROM N.A.
MEDLINE=88115604; PubMed=3429489;
Meeker T.C., Nagarajan L., Ar-Rushdi A., Croce C.M.;
"Cloning and characterization of the human PIM-1 gene: a putative
oncogene related to the protein kinases";
J. Cell. Biochem. 35:105-112(1987).
[5]
SEQUENCE OF 1-202 FROM N.A.
MEDLINE=21354098; PubMed=11460166;
Pasqualucci L., Numeister P., Goossens T., Nanjangud G.,
Chaganti R.S.K., Koppers R., Dalla-Favera R.;
"Hypermutation of multiple proto-oncogenes in B-cell diffuse
large-cell lymphomas";
Nature 412:341-346(2001).
-1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
-1- TISSUE SPECIFICITY: EXPRESSED PRIMARILY IN CELLS OF THE
HEMATOPOIETIC AND GERM LINE LINEAGES.
-1- PTM: Autophosphorylated (By similarity).
-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
PIM SUBFAMILY.
-1- DATABASE: NAME-Atlas Genet. Cytoenet. Oncol. Haematol.;
WWW="http://www.infobiogen.fr/services/chronocancer/Genes/PIMID261.html".

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M27903; AAA60090.1; -
EMBL; M16750; AAA60089.1; -
EMBL; M54915; AAA36447.1; -
EMBL; M24779; AAA81553.1; -
EMBL; AF386792; AAK70871.1; -
PIR; A27476; TVHUPL
PIR; JU0327; JU0327.
PIR; A46554; A46554.
Gene; HGNC:8986; PIM1.
MIM; 164960; -
InterPro: IPR000719; Euk_pkinase.
InterPro: IPR002290; Ser_thr_pkinase.
Pfam: PF00069; pkinase; 1.
ProDom: PD000001; Euk_pkinase; 1.
ProSITE; PS00107; PROTEIN_KINASE_ATP; 1.
ProSITE; PS00108; PROTEIN_KINASE_ST; 1.
ProSITE; PS0011; PROTEIN_KINASE_DOM; 1.
Transferase; Serine/threonine-protein kinase; Proto-oncogene;
ATP-binding; Phosphorylation.
DOMAIN 38 290 PROTEIN KINASE.
NP_BIND 44 52 ATP (BY SIMILARITY).
BINDING 67 67 ATP (BY SIMILARITY).
ACT_SITE 167 167 BY SIMILARITY.
CONFLICT 15 16 AP -> RA (IN REF. 2).
SEQUENCE 313 AA; 35685 MW; 35BA76D3668E69A3 CRC64;
Query Match 7.4%; Score 81; DB 1; Length 313;
Best Local Similarity 25.2%; Pred. No. 6.8;
Matches 39; Conservative 29; Mismatches 55; Indels 32; Gaps 9;
Y 31 REPKEIPVDKLT-----LGVISWRLNPDW---ENCENLKRREARGSYVDICDVC 81
b 85 RVPMEVLLKKVSSGSGVIRLDWFERPDSFVLLERPEPVQ-----DLDFDI 133
Y 82 PEKLPNYETIKSFEEHLHTDEIRYCLE-GSGYFDVRDQNDQWIRIALKKG--GMIVL 138
b 134 TERGALQEELARSEFFWQVL--EAVRHCHNCVGLHRDIKDN---ILIDLNRGELKLIDF 167
Y 139 PAGMYHRTDITDNYIKAMRLFVGDPVWTPYNRPH 173
Y 139 PAGMYHRTDITDNYIKAMRLFVGDPVWTPYNRPH 173

Db 188 GSGALLKDVYTD--FDGTRVY-SPPEWIRYHRYH 219
RESULT 15
PIM1_RAT
ID PIM1_RAT STANDARD; PRT; 313 AA.
AC P26794;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proto-oncogene serine/threonine-protein kinase pim-1 (EC 2.7.1.37).
GN PIM1 OR PIM-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Testis;
RX MEDLINE=92319652; PubMed=1620615;
RA Wingett D., Reeves R., Magnuson N.S.;
RT "Characterization of the testes-specific pim-1 transcript in rat.";
RL Nucleic Acids Res. 20:3183-3189(1992).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- PTM: Autophosphorylated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
PIM SUBFAMILY.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X63675; CAA45214.1; -
PIR; S26298; S26298.
InterPro: IPR000719; Euk_pkinase.
InterPro: IPR002290; Ser_thr_pkinase.
Pfam: PF00069; pkinase; 1.
ProDom: PD000001; Euk_pkinase; 1.
ProSITE; PS00107; PROTEIN_KINASE_ATP; 1.
ProSITE; PS00108; PROTEIN_KINASE_ST; 1.
ProSITE; PS0011; PROTEIN_KINASE_DOM; 1.
Transferase; Serine/threonine-protein kinase; Proto-oncogene;
ATP-binding; Phosphorylation.
DOMAIN 38 290 PROTEIN KINASE.
NP_BIND 44 52 ATP (BY SIMILARITY).
BINDING 67 67 ATP (BY SIMILARITY).
ACT_SITE 167 167 BY SIMILARITY.
SEQUENCE 313 AA; 35630 MW; D5757DA9F1821BF9 CRC64;
Query Match 7.4%; Score 81; DB 1; Length 313;
Best Local Similarity 25.2%; Pred. No. 6.8;
Matches 39; Conservative 29; Mismatches 55; Indels 32; Gaps 9;
Y 31 REPKEIPVDKLT-----LGVISWRLNPDW---ENCENLKRREARGSYVDICDVC 81
Db 85 RVPMEVLLKKVSSGSGVIRLDWFERPDSFVLLERPEPVQ-----DLDFDI 133
Y 82 PEKLPNYETIKSFEEHLHTDEIRYCLE-GSGYFDVRDQNDQWIRIALKKG--GMIVL 138
b 134 TERGALQEELARSEFFWQVL--EAVRHCHNCVGLHRDIKDN---ILIDLNRGELKLIDF 167
Y 139 PAGMYHRTDITDNYIKAMRLFVGDPVWTPYNRPH 173
Db 188 GSGALLKDVYTD--FDGTRVY-SPPEWIRYHRYH 219
Search completed: November 25, 2002, 10:09:17
Job time : 15 secs

40

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 25, 2002, 10:07:25 ; Search time 31 Seconds
(without alignments)
1322.690 Million cell updates/sec

Title: US-09-785-738a-2

Perfect score: 1097
Sequence: 1 MENEFQDGKTEVIEAWYMD.....KEFLAKLLKSEGENAQAVEGF 199

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 205047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

(SPTREMBL_21.4)

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organalle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1097	100.0	199	10	O65035 oryza sativ
2	955.5	87.1	198	10	O81276 oryza sativ
3	946.5	86.3	198	10	Q93XJ5 hordeum vul
4	833.5	76.0	199	10	Q8W108 arabidopsis
5	789.5	72.0	940	10	O23327 arabidopsis
6	781	71.2	199	10	O48707 arabidopsis
7	680	62.0	185	10	Q9FG79 arabidopsis
8	609	55.5	179	11	Q99J79 mus musculus
9	586	53.4	179	4	Q9BV57 homo sapien
10	581	53.0	179	4	Q9NV57 homo sapien
11	451	41.1	342	5	O9VTB1 drosophila
12	432	39.4	116	4	O9GP85 homo sapien
13	307.5	28.0	178	3	O94286 schizosacch
14	307.5	28.0	178	5	Q20340 caenorhabdi
15	287	26.2	71	10	Q93W70 musa acumin
16	229	20.9	221	5	Q09407 caenorhabdi

P91416 caenorhabdi
O31669 bacillus su
O61597 osteria
Q9bda4 xylella fas
O67785 aquifex aeo
Q91341 pseudomonas
Q8vpv0 synectococc
Q8ytj3 anabaena sp
Q9zfe7 klebsiella
Q9rm34 gluconobact
Q936j4 rhodospiril
Q18910 caenorhabdi
Q92t91 rhizobium m
Q99xq5 streptococc
O8t154 methanosarc
Q8tyw3 methanopyru
Q9skv9 arabidopsis
Q929t1 bacillus ha
O26046 helicobacte
Q08837 triticum ae
Q9f7g3 salmonella
Q9vq73 drosophila
P91740 hydra atten
Q8sv03 encephalito
Q9vx09 drosophila
Q9r0x5 mus musculu
Q97hf6 clostridium
Q9f7g4 salmonella
O66618 aquifex aeo

ALIGNMENTS

RESULT 1

O65035 PRELIMINARY; PRT; 199 AA.
AC O65035
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Submergence induced protein 2.
GN SIP2.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
CX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Lorbietcke R., Sauter M.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF050200; AAC05511.1; -;
DR InterPro; IPR004313; Acired_dioxygenase.
DR Pfam; PF03079; ARD; 1.
SQ SEQUENCE 199 AA; 23601 MW; D1938253A1FC954D CRC64;

Query Match 100.0%; Score 1097; DB 10; Length 199;
Best Local Similarity 100.0%; Pred. No. 6.4e-94;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MENEFQDGKTEVIEAWYMDSEEDORLPHPHREPKETIPVDKLTGELGVISWRLNPDNWC 60
Db 1 MENEFQDGKTEVIEAWYMDSEEDORLPHPHREPKETIPVDKLTGELGVISWRLNPDNWC 60
Qy 61 ENLKRIEARGYSYVDICDVCPEKLPNYETIKISFFEEHLHTDEEIRYCLEGSGYFDVRD 120
Db 61 ENLKRIEARGYSYVDICDVCPEKLPNYETIKISFFEEHLHTDEEIRYCLEGSGYFDVRD 120
Qy 121 QNDQWIRIALKKGNIIVLPAGMYHRFTLTDNYIKAMRLFVGPDPVWTPYNRPHDHPARK 180
Db 121 QNDQWIRIALKKGNIIVLPAGMYHRFTLTDNYIKAMRLFVGPDPVWTPYNRPHDHPARK 180

```
QY 181 EFLAKLLKSEGENQAVEGF 199
|||||
Db 181 EFLAKLLKSEGENQAVEGF 199
|||||

RESULT 2
O81276 PRELIMINARY; PRT; 198 AA.
AC O81276;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Submergence induced protein 2A.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Lorbiecke R., Sauter M.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF068332; AAC19375.1; -
DR InterPro; IPR004313; Acired_dioxygenase.
DR Pfam; PF03079; ARD; 1.
SQ SEQUENCE 198 AA; 23498 MW; A7165C7A7B80E92E CRC64;

Query Match 87.1%; Score 955.5; DB 10; Length 198;
Best Local Similarity 85.3%; Pred. No. 8.3e-81;
Matches 168; Conservative 19; Mismatches 9; Indels 1; Gaps 1;

QY 1 MENEFDGKTEVIEAWYMDSEEDQRLPHHREPKETIPVDKLTGELGVISWRLNPDNWNENC 60
|||||
Db 1 MENEFDGKTEVIEAWYMDSEEDQRLPHHREPKETIPVDKLTGELGVISWRLNPDNWNENC 60
|||||

QY 61 ENLKRIREARGSYVDICDVCPEKLPNYETKIKSFEEHLHTDEEIRYCLEGSGYFDVRD 120
|||||
Db 61 ENLKRIREARGSYVDICDVCPEKLPNYETKIKSFEEHLHTDEEIRYCLEGSGYFDVRD 120
|||||

QY 121 QNDQWIRIATKKGGMIVLPAGMYHRTLTDTNVIKAMRLFVGDVPWTPYNNRPHDHLPAK 180
|||||
Db 121 QNDQWIRIATKKGGMIVLPAGMYHRTLTDTNVIKAMRLFVGDVPWTPYNNRPHDHLPAK 180
|||||

QY 181 EFLAKLLKSEGENQAVE 197
|||||
Db 181 EYVEKII-NRGGTQAVE 196
|||||

RESULT 3
Q93XJ5 PRELIMINARY; PRT; 198 AA.
AC Q93XJ5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ID11 protein.
DE ID11 protein.
GN ID11.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RA Yamaguchi H., Nakanishi H., Nishizawa N.K., Mori S.;
RT "Induction of the ID11 gene in Fe-deficient barley roots: a gene
encoding a putative enzyme that catalyses the methionine salvage
pathway for phytosiderophore production.";
RL Soil Sci. Plant Nutrition 46:1-9(2000).
DR EMBL; AB025597; BAB61039.1; -
DR InterPro; IPR004313; Acired_dioxygenase.
DR Pfam; PF03079; ARD; 1.

QY 181 EFLAKLLKSEGENQAVEGF 199
|||||
Db 181 EYVEKII-NRGGTQAVE 196
|||||

RESULT 3
Q93XJ5 PRELIMINARY; PRT; 198 AA.
AC Q93XJ5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ID11 protein.
GN ID11.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RA Yamaguchi H., Nakanishi H., Nishizawa N.K., Mori S.;
RT "Induction of the ID11 gene in Fe-deficient barley roots: a gene
encoding a putative enzyme that catalyses the methionine salvage
pathway for phytosiderophore production.";
RL Soil Sci. Plant Nutrition 46:1-9(2000).
DR EMBL; AB025597; BAB61039.1; -
DR InterPro; IPR004313; Acired_dioxygenase.
DR Pfam; PF03079; ARD; 1.

QY 181 EFLAKLLKSEGENQAVEGF 199
|||||
Db 181 EYVEKII-NRGGTQAVE 196
|||||

RESULT 4
Q8W108 PRELIMINARY; PRT; 199 AA.
AC Q8W108;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE A74914710/dl3395c.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Banh J., Bowser L.,
RA Carinci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF462818; AAL58908.1; -
DR InterPro; IPR004313; Acired_dioxygenase.
DR Pfam; PF03079; ARD; 1.
SQ SEQUENCE 199 AA; 23356 MW; 3FD427B218AF80A4 CRC64;

Query Match 76.0%; Score 833.5; DB 10; Length 199;
Best Local Similarity 76.5%; Pred. No. 1.7e-69;
Matches 143; Conservative 29; Mismatches 14; Indels 1; Gaps 1;

QY 6 ODGKTEVIEAWYMDSEEDQRLPHHREPKETIPVDKLTGELGVISWRLNPDNWNENC 65
|||||
Db 6 ODGKTEVIEAWYMDSEEDQRLPHHREPKETIPVDKLTGELGVISWRLNPDNWNENC 65
|||||

QY 66 TREARGSYVDICDVCPEKLPNYETKIKSFEEHLHTDEEIRYCLEGSGYFDVRDNDOW 125
|||||
Db 66 TREARGSYVDICDVCPEKLPNYETKIKSFEEHLHTDEEIRYCLEGSGYFDVRDNDOW 125
|||||

QY 126 TRIALKKGGMIVLPAGMYHRTLTDTNVIKAMRLFVGDVPWTPYNNRPHDHLPAK 185
|||||
Db 126 IRVWYKKGGMIVLPAGMYHRTLTDTNVIKAMRLFVGDVPWTPYNNRPHDHLPAK 185
|||||

QY 186 LIK-SEG 191
|||||
Db 186 FVKVNEG 192
|||||
```



```

RESULT 5
O23327 PRELIMINARY; PRT; 940 AA.
AC O23327
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 108.0 kDa protein.
GN ATAG14710
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terry N.,
RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
RA Pulgomech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
RA Jones J., Palme K., Anson W., Delsen M., Bancroft I., Mewes H.W.,
RA Schueller C., Chalwatzis N.,
RA Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; 297336; CAB10250.1; -.
DR EMBL; AL161539; CAB78513.1; -.
DR InterPro; IPR004313; Acired_dioxygenase.
DR Pfam; PF03079; ARD; 2.
KW Hypothetical protein.
SQ SEQUENCE 940 AA; 107963 MW; 0451A8C9C4371179 CRC64;

Query Match 72.0%; Score 789.5; DB 10; Length 940;
Best Local Similarity 66.3%; Pred. No. 1.4e-64;
Matches 138; Conservative 26; Mismatches 15; Indels 29; Gaps 1;

QY 1 MENEFDGKTEVIAWYMDSEEDQRLPHHREPKEFIPVDKLTGVLISWRNPNWENC 60
Db 1 MGEAVKQREVIQAWYMDSEEDQRLPHHREPKEFIPVDKLTGVLISWRNPNWENC 60
QY 61 ENLKIRAEARGSYV-----DIDVCCEKLPNYETK 91
Db 61 EDLKKIRSRGYSYVWCGLYFTIPLDLNGRCLFLSKSYFNSLGDCEVCCEKLPNYEVK 120
QY 92 IKSFFEEHLHTDEIRYCLESGYFVDRQNDQWIRALKKGGMIVLPAGMYHRTD 151
Db 121 VKSFFEEHLHTDEIRYCVAGTYGTVDRDRNEAIRVLVKKGGMIVLPAGTYHRTVDS 180
QY 152 NYIKAMRLFVGDVWTPYNRPHDLPAR 179
Db 181 NYIKAMRLFVGEVWTPYNRPHDLPAR 208

RESULT 6
O48707 PRELIMINARY; PRT; 199 AA.
AC O48707
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE At2g26400 protein.
GN AT2G26400.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,

RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.-J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome II of Arabidopsis thaliana.";
RL Nature 402:761-768(1999).
DR EMBL; AC002505; AAC14490.1; -.
DR InterPro; IPR004313; Acired_dioxygenase.
DR Pfam; PF03079; ARD; 1.
SQ SEQUENCE 199 AA; 23441 MW; 73053F8BC4ADEE25 CRC64;

Query Match 71.2%; Score 781; DB 10; Length 199;
Best Local Similarity 71.7%; Pred. No. 1.2e-64;
Matches 134; Conservative 27; Mismatches 26; Indels 0; Gaps 0;

QY 6 ODGKTEVIAWYMDSEEDQRLPHHREPKEFIPVDKLTGVLISWRNPNWENC 65
Db 6 KDQTEVIAWYMDSEEDQRLPHHREPKEFIPVDKLTGVLISWRNPNWENC 65
QY 66 IREARGSYVDICDVCCEKLPNYETKIKSFEEHLHTDEIRYCLESGYFVDRQNDQ 125
Db 66 IRESRGYSYMDLCEVCCEKLPNYEKKVMEFEEHLHTDEIRYCLESGYFVDRQNDQ 125
QY 126 IRTALKKGGMIVLPAGMYHRTDTONYIKAMRLFVGDVWTPYNRPHDLPARKEFLAK 185
Db 126 IRTVKKKGLIVFPAGTYHRTVDSNDYMKAMRLFVGGPVMTAYNRPHDLPARKAYMKK 185
QY 186 LKSEGE 192
Db 186 FLKVID 192

RESULT 7
Q9FG79 PRELIMINARY; PRT; 185 AA.
AC Q9FG79
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Genomic DNA, chromosome 5, pl clone:MOB19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB026651; BAB11314.1; -.
DR InterPro; IPR004313; Acired_dioxygenase.
DR Pfam; PF03079; ARD; 1.
SQ SEQUENCE 185 AA; 22263 MW; A4A4F594ED6BC2DA CRC64;

Query Match 62.0%; Score 680; DB 10; Length 185;
Best Local Similarity 64.4%; Pred. No. 2.6e-55;
Matches 116; Conservative 29; Mismatches 35; Indels 0; Gaps 0;

QY 13 IEAWYMDSEEDQRLPHHREPKEFIPVDKLTGVLISWRNPNWENC 72
Db 1 MQAWFMDSDNEQRLPHHREPKEFIPVDKLTGVLISWRNPNWENC 60
QY 73 SYVDICDVCCEKLPNYETKIKSFEEHLHTDEIRYCLESGYFVDRQNDQWIRALKK 132
Db 61 DYMDLLDLCPEKVSNEYEKLNFETEIHRDEIRYCLESGYFVDRQNDQWIRALKK 120
QY 133 GGMIVLPAGMYHRTDTONYIKAMRLFVGDVWTPYNRPHDLPARKEFLAKLSEGE 192

```

```
Db 121 GOLIVLPAGIYHRTLDASNYIKMLRFLVGPVWTPYNNRQBEHPVRKKYIHGLTYKFE 180
RESULT 8
Q99J79
ID Q99J79 PRELIMINARY; PRT; 179 AA.
AC Q99J79;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to hypothetical protein FLJ10913.
GN AL024210.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005695; AAH05695.1; -.
DR MGI; MGI:2144929; AL024210.
DR InterPro; IPR004313; Acired_dioxgnase.
DR Pfam; PF03079; ARD; 1.
KW Hypothetical protein.
SQ SEQUENCE 179 AA; 21523 MW; 02CE7B89181EBEE CRC64;
Query Match 55.5%; Score 609; DB 11; Length 179;
Best Local Similarity 60.0%; Pred. No. 9.5e-49;
Matches 105; Conservative 32; Mismatches 38; Indels 0; Gaps 0;
QY 12 VIEAWYMDSEEDQRLPHHREPKFIPVDKLTGLGVISWRLNPDNWNENCKLRIRREARG 71
Db 1 MVQAWYMDADPGDPQHPRPDPGRPVGLQLRRLGLVLYWKLADKYENDPELEKIRREARN 60
QY 72 YSYVDICDVCEKLPNYETKIKSFEEHLHTDEIRYCLGSGYFDVRDQNDQWIRIATK 131
Db 61 YSWMDIITICKDKLPNYEEKIKMFEYEHLDDEIRYILDSGYFDVRDQNDQWIRIFME 120
QY 132 KGMIVLPAGMYHRTLDTONYIKAMRLFVGDVWTPYNNRPHDHLPAKKEFLAKLLKS 189
Db 121 KGMVTLUPAGIYHRTVDEKNTYKAMRLFVGPVWTPYNNRPHDHLPAKKEFLAQT 178
RESULT 10
Q9NV57
ID Q9NV57 PRELIMINARY; PRT; 179 AA.
AC Q9NV57;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE OVARC100209 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA Masuho Y., Kanehori K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001775; BAA91901.1; -.
DR InterPro; IPR004313; Acired_dioxgnase.
DR Pfam; PF03079; ARD; 1.
SQ SEQUENCE 179 AA; 21483 MW; 3B67E975EAD68A1 CRC64;
Query Match 53.0%; Score 581; DB 4; Length 179;
Best Local Similarity 56.7%; Pred. No. 3.8e-46;
Matches 101; Conservative 33; Mismatches 44; Indels 0; Gaps 0;
QY 12 VIEAWYMDSEEDQRLPHHREPKFIPVDKLTGLGVISWRLNPDNWNENCKLRIRREARG 71
Db 1 MVQAWYMDADPGDPQHPRPDPGRPVGLQLRRLGLVLYWKLADKYENDPELEKIRREARN 60
QY 72 YSYVDICDVCEKLPNYETKIKSFEEHLHTDEIRYCLGSGYFDVRDQNDQWIRIATK 131
Db 61 YSWMDIITICKDKLPNYEEKIKMFEYEHLDDEIRYILDSGYFDVRDQNDQWIRIFME 120
QY 132 KGMIVLPAGMYHRTLDTONYIKAMRLFVGDVWTPYNNRPHDHLPAKKEFLAKLLKS 189
Db 121 KGMVTLUPAGIYHRTVDEKNTYKAMRLFVGPVWTPYNNRPHDHLPAKKEFLAQT 178
RESULT 11
Q9VTB1
ID Q9VTB1 PRELIMINARY; PRT; 342 AA.
AC Q9VTB1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG5000 protein.
GN CG7983 OR CG8000.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
```

```
Db 121 GOLIVLPAGIYHRTLDASNYIKMLRFLVGPVWTPYNNRQBEHPVRKKYIHGLTYKFE 180
RESULT 8
Q99J79
ID Q99J79 PRELIMINARY; PRT; 179 AA.
AC Q99J79;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to hypothetical protein FLJ10913.
GN AL024210.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005695; AAH05695.1; -.
DR MGI; MGI:2144929; AL024210.
DR InterPro; IPR004313; Acired_dioxgnase.
DR Pfam; PF03079; ARD; 1.
KW Hypothetical protein.
SQ SEQUENCE 179 AA; 21523 MW; 02CE7B89181EBEE CRC64;
Query Match 55.5%; Score 609; DB 11; Length 179;
Best Local Similarity 60.0%; Pred. No. 9.5e-49;
Matches 105; Conservative 32; Mismatches 38; Indels 0; Gaps 0;
QY 12 VIEAWYMDSEEDQRLPHHREPKFIPVDKLTGLGVISWRLNPDNWNENCKLRIRREARG 71
Db 1 MVQAWYMDADPGDPQHPRPDPGRPVGLQLRRLGLVLYWKLADKYENDPELEKIRREARN 60
QY 72 YSYVDICDVCEKLPNYETKIKSFEEHLHTDEIRYCLGSGYFDVRDQNDQWIRIATK 131
Db 61 YSWMDIITICKDKLPNYEEKIKMFEYEHLDDEIRYILDSGYFDVRDQNDQWIRIFME 120
QY 132 KGMIVLPAGMYHRTLDTONYIKAMRLFVGDVWTPYNNRPHDHLPAKKEFLAKLLKS 189
Db 121 KGMVTLUPAGIYHRTVDEKNTYKAMRLFVGPVWTPYNNRPHDHLPAKKEFLAQT 178
RESULT 10
Q9NV57
ID Q9NV57 PRELIMINARY; PRT; 179 AA.
AC Q9NV57;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE OVARC100209 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,
RA Masuho Y., Kanehori K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001775; BAA91901.1; -.
DR InterPro; IPR004313; Acired_dioxgnase.
DR Pfam; PF03079; ARD; 1.
SQ SEQUENCE 179 AA; 21483 MW; 3B67E975EAD68A1 CRC64;
Query Match 53.0%; Score 581; DB 4; Length 179;
Best Local Similarity 56.7%; Pred. No. 3.8e-46;
Matches 101; Conservative 33; Mismatches 44; Indels 0; Gaps 0;
QY 12 VIEAWYMDSEEDQRLPHHREPKFIPVDKLTGLGVISWRLNPDNWNENCKLRIRREARG 71
Db 1 MVQAWYMDADPGDPQHPRPDPGRPVGLQLRRLGLVLYWKLADKYENDPELEKIRREARN 60
QY 72 YSYVDICDVCEKLPNYETKIKSFEEHLHTDEIRYCLGSGYFDVRDQNDQWIRIATK 131
Db 61 YSWMDIITICKDKLPNYEEKIKMFEYEHLDDEIRYILDSGYFDVRDQNDQWIRIFME 120
QY 132 KGMIVLPAGMYHRTLDTONYIKAMRLFVGDVWTPYNNRPHDHLPAKKEFLAKLLKS 189
Db 121 KGMVTLUPAGIYHRTVDEKNTYKAMRLFVGPVWTPYNNRPHDHLPAKKEFLAQT 178
RESULT 11
Q9VTB1
ID Q9VTB1 PRELIMINARY; PRT; 342 AA.
AC Q9VTB1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG5000 protein.
GN CG7983 OR CG8000.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
```

```
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasako P., Lei Y., Levitsky A.A., Li J., Li Z., Lian Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003547; AA050141.1;
DR Flybase: FB90036036; CG7983.
DR InterPro: IPR004313; Acired_dioxygenase.
DR Pfam: PF03079; ARD; 1.
SQ SEQUENCE 342 AA; 40236 MW; D6A587C4D1E3F67C CRC64;
```

```
Query Match 41.1%; Score 451; DB 5; Length 342;
Best Local Similarity 43.5%; Pred. No. 9.5e-34;
Matches 84; Conservative 30; Mismatches 47; Indels 32; Gaps 3;
```

```
QY 12 VIEAWYMDSEEDRLPHRPEKFIIPVD-----KLTELGVISWLNPDNWNENC 60
Db 1 MVQVWYMDTEEDQRLHHRNPAYLELDLYQKTGYEYFKATYLDHSHLQINADYQSD 60
QY 61 ENLKIRARGSYVDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLEGSGYFVD 120
Db 61 NTLTEL-----AKHYANKLKAFTEHLHTDEIRLILEGSGYFVD 103
QY 121 QNDQWIRIALKGGMVLVPGMYHRTLDTONYIKAMRLFVGDVPWTPYRPHDHPARK 180
Db 104 NEDNWLRLKVVKGDLIIIPAGIYHRTLDTONYIRYFVGPVWPAHNPADMDCKR 163
QY 181 EFLAKLKEGEN 193
Db 164 SY-----IKHQSEN 172
```

```
RESULT 12
Q96P85
ID Q96P85 PRELIMINARY; PRT; 116 AA.
AC Q96P85;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE S1PL.
```

```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21498672; PubMed=11602742;
RA Yeh C.T., Lai H.Y., Chen T.C., Chu C.M., Liaw Y.F.;
RT "Identification of a hepatic factor capable of supporting hepatitis C
RT virus replication in a nonpermissive cell line.";
RL J. Virol. 75:11017-11024(2001).
DR EMBL: AF403478; AAL25800.1;
DR InterPro: IPR004313; Acired_dioxygenase.
DR Pfam: PF03079; ARD; 1.
SQ SEQUENCE 116 AA; 13892 MW; 3C26CC366949E3B6 CRC64;
```

```
Query Match 39.4%; Score 432; DB 4; Length 116;
Best Local Similarity 65.2%; Pred. No. 1.4e-32;
Matches 75; Conservative 18; Mismatches 22; Indels 0; Gaps 0;
```

```
QY 75 VDICDVCPEKLPNYETKIKSFEEHLHTDEIRYCLEGSGYFVDQNDQWIRIALKGG 134
Db 1 MDIITCKDLPNVEEIKMFYEEHLHLDDEIRYLDGSGYFVDKEDQWIRIFMEKGD 60
QY 135 MIVLPAGMYHRTLDTONYIKAMRLFVGDVPWTPYRPHDHPARKEFLAKLKS 189
Db 61 MVTLPAGIYHRTVDKKNYTKAMRLFVGPVWTAYNRPADHFEARGQYVKFLAQ 115
```

```
RESULT 13
O94286
ID O94286 PRELIMINARY; PRT; 178 AA.
AC O94286;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 20.9 kDa protein.
GN SP6C887.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Lyne M., Rajandream M.A., Barrell B.G., Lelaure V., Galibert F.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL033388; CAA21886.1;
DR InterPro: IPR004313; Acired_dioxygenase.
DR Pfam: PF03079; ARD; 1.
KW Hypothetical protein.
SQ SEQUENCE 178 AA; 20936 MW; 73DDAA2874371D90 CRC64;
```

```
Query Match 28.0%; Score 307.5; DB 3; Length 178;
Best Local Similarity 40.1%; Pred. No. 8.4e-21;
Matches 71; Conservative 32; Mismatches 65; Indels 9; Gaps 5;
```

```
QY 17 YMDSEEDRLPHRPEKFIIPVDKLTGVLISWLNPDNWNENCNKRIRARGSYVD 76
Db 4 YIFQDEGDQKPN--DSKTEVSAEDL-EAAKVSYRHHGDGLHTFAD--GLMKEYGFKNRD 58
QY 77 ICDCVCEKL-PNYETKIKSFEEHLHTDEIRYCLEGSGYFVDQNDQWIRIALKGGM 135
Db 59 EVVVSRLGGLGDYDNMVKKFFEEHLHDEDEIRLIDGNGYFVRSVDDRWVRFVEKGL 118
QY 136 IVLPGMYHRTLDTONYIKAMRLFVGDVPWTPYRPHDHPARKEFLAKLKS 189
Db 119 IILPGIYHRTTTDDYIHAMRPFHENPKWIALSRDSTSELDARKSYLNSIKKS 175
```

```
RESULT 14
Q20340
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 18:50:13 ; Search time 3125 Seconds
(without alignments)
8120.847 Million cell updates/sec

Title: US-09-785-738A-1
Perfect score: 872
Sequence: 1 agacgaacaaacacagaat.....acaggagctctgaagctgaa 872

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBank

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pi.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	872	100.0	872	6	AX277504	AX277504 Sequence
2	872	100.0	872	8	AF050200	AF050200 Oryza sat
3	424.4	48.7	980	6	AX277506	AX277506 Sequence
4	424.4	48.7	980	8	AF068332	AF068332 Oryza sat
5	417.2	47.8	997	8	AB025597	AB025597 Hordeum v
6	339.2	38.9	933	6	AX277514	AX277514 Sequence
7	312.2	35.8	840	8	AF462818	AF462818 Arabidops
8	311.8	35.8	774	6	AX277508	AX277508 Sequence
9	310.6	35.6	991	8	AY086754	AY086754 Arabidops
10	304.8	35.0	1977	6	AX449332	AX449332 Sequence
11	303.2	34.8	925	8	AY089014	AY089014 Arabidops
12	285.6	32.8	889	6	AX277512	AX277512 Sequence
13	255.6	29.3	142737	8	AC027658	AC027658 Oryza sat
14	244	28.0	759	8	AY086650	AY086650 Arabidops
15	242	27.8	603	6	AX277510	AX277510 Sequence
16	209.4	24.0	1685	10	BC005695	BC005695 Mus muscu
17	207.8	23.8	972	6	AX277518	AX277518 Sequence
18	201.4	23.1	1123	9	BC001467	BC001467 Homo sapi
19	199.8	22.9	624	6	AX449303	AX449303 Sequence
20	199.8	22.9	1107	6	AR116047	AR116047 Sequence
21	198.2	22.7	1628	9	AK001775	AK001775 Homo sapi
22	182.4	20.9	706	6	AX277520	AX277520 Sequence
23	181.8	20.8	920	6	AX277516	AX277516 Sequence
24	172	19.7	1093	6	E32094	E32094 Human gastr
25	167	19.2	559	9	AF403478	AF403478 Homo sapi
26	159.4	18.3	286	11	G71630	G71630 AG322334FM
27	154.8	17.8	456	6	AX408702	AX408702 Sequence
28	151.2	17.3	308	11	G70864	G70864 A6372334FB
29	136.8	15.7	300	6	AX429573	AX429573 Sequence
30	116	13.3	87286	8	AB026651	AB026651 Arabidops
31	110.6	12.7	532	8	AF414129	AF414129 Musa acum
32	104.6	12.0	44308	8	SPBC887	AL033388 S.pombe c
33	102.2	11.7	177327	8	AC105729	AC105729 Oryza sat
34	97.6	11.2	141054	2	AC109110	AC109110 Rattus no
35	97.6	11.2	169103	2	AC125638	AC125638 Rattus no
36	96.2	11.0	480	8	CNS01AHT	AL112953 Botrytis
37	91.2	10.5	29163	9	HSJ450M14	AL132654 Human DNA
38	89.6	10.3	119249	9	AC114810	AC114810 Homo sapi
39	86	9.9	113457	9	AC003686	AC003686 Homo sapi
40	86	9.9	168085	9	AC026222	AC026222 Homo sapi
41	86	9.9	278581	2	AC074326	AC074326 Homo sapi
42	79.2	9.1	197976	8	ATCHRIV39	AL161539 Arabidops
43	79.2	9.1	206606	8	ATFCAL	297336 Arabidops
44	78.6	9.0	131905	2	OSJN00131	AL606995 Oryza sat
45	78.2	9.0	32481	8	SC8270	Z48613 S.cerevisia

ALIGNMENTS

RESULT 1	AX277504	AX277504	872 bp	DNA	linear	PAT 29-OCT-2001
LOCUS	Sequence 1 from Patent WO0161020.					
DEFINITION	AX277504					
ACCESSION	AX277504					
VERSION	AX277504.1	GI:16549009				
KEYWORDS	Oryza sativa.					
SOURCE	Oryza sativa					
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;					
	Ehrhartoideae; Oryzaceae; Oryza.					
REFERENCE	1					
AUTHORS	Sauter,M.M. and Lorbiecke,R.					
TITLE	Alteration of growth and adaptation under hypoxic conditions					

JOURNAL Patent: WO 0161020-A 1 23-AUG-2001;
CropDesign N.V. (BE)
FEATURES Location/Qualifiers
source 1..872
/organism="Oryza sativa"
/db_xref="taxon:4530"
69..668
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10342.1"
/db_xref="GI:16549010"
/translation="MENEFODGKTEVIAWYMDSEODRLPHHREPKFIPVDKLTE
LGVISWRLNPDNWNENLKRIRIARGYSYVDICDVCPEKLPNVETKIKSFFREHLHT
DEIRYCLEGSYFDVRDNDQNDORIRALKGGMVLPAGMYHRTLDTDNYIKAMRLF
VGDPVWTPYINRPHDLPARKEFLAKLLKSEGENQAVEGF"
BASE COUNT 267 a 176 c 206 g 223 t
ORIGIN
Query Match 100.0%; Score 872; DB 6; Length 872;
Best Local Similarity 100.0%; Pred. No. 8e-235;
Matches 872; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGACGACAAAAACAGAGATCCATCGCCATAATCGAAGGTTTCGCTTCCACCCCG 60
DB 1 AGACGACAAAAACAGAGATCCATCGCCATAATCGAAGGTTTCGCTTCCACCCCG 60
QY 61 CAATCCACATGGAGAACGAATCCAGAGTGGTAAGACGAGGTGATAGAGCATGGTACA 120
DB 61 CAATCCACATGGAGAACGAATCCAGAGTGGTAAGACGAGGTGATAGAGCATGGTACA 120
QY 121 TGGATGATACGAGAGGACGAGAGGCTTCCTCATCCGCGAACCCAAAGAAATTCATTC 180
DB 121 TGGATGATACGAGAGGACGAGAGGCTTCCTCATCCGCGAACCCAAAGAAATTCATTC 180
QY 181 CTGTTGATAAGCTTACAGAACTAGAGTAAATCAGCTGGCGCCTAAATCTGTAACTGGG 240
DB 181 CTGTTGATAAGCTTACAGAACTAGAGTAAATCAGCTGGCGCCTAAATCTGTAACTGGG 240
QY 181 CTGTTGATAAGCTTACAGAACTAGAGTAAATCAGCTGGCGCCTAAATCTGTAACTGGG 240
DB 181 CTGTTGATAAGCTTACAGAACTAGAGTAAATCAGCTGGCGCCTAAATCTGTAACTGGG 240
QY 241 AGAATTGGCAGAACCTGAAGAAATCCCGGAGCCAGAGGTTACTCTTATGTGGACATTT 300
DB 241 AGAATTGGCAGAACCTGAAGAAATCCCGGAGCCAGAGGTTACTCTTATGTGGACATTT 300

QY 781 AAGATTAAATGATTTCGTTAAATCTTATGATCAATCGCATATAGCATTTCCGAATG 840
DB 781 AAGATTAAATGATTTCGTTAAATCTTATGATCAATCGCATATAGCATTTCCGAATG 840
QY 841 TGTTCCTCAATAAACAGAGGATCATGAAGCTGAA 872
DB 841 TGTTCCTCAATAAACAGAGGATCATGAAGCTGAA 872
RESULT 2
AF050200 872 bp mRNA linear PLN 12-MAR-1998
LOCUS Oryza sativa submergence induced protein 2 (sip2) mRNA, complete cds.
DEFINITION
ACCESSION AF050200 GI:2952337
VERSION AF050200
KEYWORDS
SOURCE
ORGANISM Oryza sativa.
Oryza sativa.
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 872)
AUTHORS Lorbiecke, R. and Sauter, M.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-1998) Department of General Botany, AMP II,
University of Hamburg, Ohnhorststrasse 18, Hamburg 22609, Germany
FEATURES
source 1..872
/organism="Oryza sativa"
/db_xref="taxon:4530"
gene 1..872
/gene="SIP2"
69..668
/gene="SIP2"
/note="similar to the Arabidopsis thaliana product encoded by the sequence presented in GenBank Accession Number AC002505"
/codon_start=1
/product="submergence induced protein 2"
/protein_id="AAC0511.1"
/db_xref="GI:2952338"
/translation="MENEFODGKTEVIAWYMDSEODRLPHHREPKFIPVDKLTE
LGVISWRLNPDNWNENLKRIRIARGYSYVDICDVCPEKLPNVETKIKSFFREHLHT
DEIRYCLEGSYFDVRDNDQNDORIRALKGGMVLPAGMYHRTLDTDNYIKAMRLF
VGDPVWTPYINRPHDLPARKEFLAKLLKSEGENQAVEGF"
BASE COUNT 267 a 176 c 206 g 223 t
ORIGIN
Query Match 100.0%; Score 872; DB 8; Length 872;
Best Local Similarity 100.0%; Pred. No. 8e-235;
Matches 872; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGACGACAAAAACAGAGATCCATCGCCATAATCGAAGGTTTCGCTTCCACCCCG 60
DB 1 AGACGACAAAAACAGAGATCCATCGCCATAATCGAAGGTTTCGCTTCCACCCCG 60
QY 61 CAATCCACATGGAGAACGAATTCAGAGTGGTAAGACGAGGTGATAGAGCATGGTACA 120
DB 61 CAATCCACATGGAGAACGAATTCAGAGTGGTAAGACGAGGTGATAGAGCATGGTACA 120
QY 121 TGGATGATACGAGAGGACGAGAGGCTTCCTCATCCGCGAACCCAAAGAAATTCATTC 180
DB 121 TGGATGATACGAGAGGACGAGAGGCTTCCTCATCCGCGAACCCAAAGAAATTCATTC 180
QY 181 CTGTTGATAAGCTTACAGAACTAGAGTAAATCAGCTGGCGCCTAAATCTGTAACTGGG 240
DB 181 CTGTTGATAAGCTTACAGAACTAGAGTAAATCAGCTGGCGCCTAAATCTGTAACTGGG 240
QY 241 AGAATTGGCAGAACCTGAAGAAATCCCGGAGCCAGAGGTTACTCTTATGTGGACATTT 300
DB 241 AGAATTGGCAGAACCTGAAGAAATCCCGGAGCCAGAGGTTACTCTTATGTGGACATTT 300

QY 301 GTGATGTGTCGCCAGAGAGCTGCCAAATATTAAGAACTAAGATCAAGAGATTCTTTGAAG 360
DB 301 GTGATGTGTCGCCAGAGAGCTGCCAAATATTAAGAACTAAGATCAAGAGATTCTTTGAAG 360
QY 361 AACACCTGCATACCGATGAAGAAATACGCTATTCTCTTGAAGGAGTGATACATTGATG 420
DB 361 AACACCTGCATACCGATGAAGAAATACGCTATTCTCTTGAAGGAGTGATACATTGATG 420
QY 421 TGAGAGACCAAAATGATCAGTGGATTCGTATAGCACTGAAGAAAGGAGGATGATGTTTC 480
DB 421 TGAGAGACCAAAATGATCAGTGGATTCGTATAGCACTGAAGAAAGGAGGATGATGTTTC 480
QY 481 TGCCTGAGGATGTACACCGCTTTAGCTTGGACACCGCACTATATCAAGGCAATGC 540
DB 481 TGCCTGAGGATGTACACCGCTTTAGCTTGGACACCGCACTATATCAAGGCAATGC 540
QY 541 GACTGTGTTGTTGGGATCCTGTTGGACACCTCAACCGTCCCATCACCATCTTCCTG 600
DB 541 GACTGTGTTGTTGGGATCCTGTTGGACACCTCAACCGTCCCATCACCATCTTCCTG 600
QY 601 CAAGAAAGGAGCTTTTGGCTAAACTCTCAAGTCAGAAAGTGAAATCAAGCAGTTGAAG 660
DB 601 CAAGAAAGGAGCTTTTGGCTAAACTCTCAAGTCAGAAAGTGAAATCAAGCAGTTGAAG 660
QY 661 GCTTCTGAGGCTTTTGTGGCTCCTGCACCTCGGTTCTATATCAACCTGAATGAAGATG 720
DB 661 GCTTCTGAGGCTTTTGTGGCTCCTGCACCTCGGTTCTATATCAACCTGAATGAAGATG 720
QY 721 TGCATAGCAATGTAAATTTAGCACAGTGGCTATGGTCCGCACTCAACCACTGAAGTGA 780
DB 721 TGCATAGCAATGTAAATTTAGCACAGTGGCTATGGTCCGCACTCAACCACTGAAGTGA 780
QY 781 AAGATTAAATGATTTCTGTTAACTTCTTATGATCAATCGCATATAGCATTTCCGAAATG 840
DB 781 AAGATTAAATGATTTCTGTTAACTTCTTATGATCAATCGCATATAGCATTTCCGAAATG 840
QY 841 TGTTCATAAAGAGGAGTCAATGAAGCTGAA 872
DB 841 TGTTCATAAAGAGGAGTCAATGAAGCTGAA 872

RESULT 3
AX277506
LOCUS AX277506 980 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 3 from Patent WO0161020.
ACCESSION AX277506
VERSION AX277506.1 GI:16549011
KEYWORDS
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
1 Sauter, M.M. and Lorbiecke, R.
TITLE Alteration of growth and adaptation under hypoxic conditions
JOURNAL Patent: WO 0161020-A 3 23-AUG-2001;
CropDesign N.V. (BE)

FEATURES
Source 1..980
/organism="Oryza sativa"
/db_xref="taxon:4530"
139..735
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD10343.1"
/db_xref="gi:16549012"
/translation="MENQFDQKEVIEAWYMDSEEDQRLPHHREPKEFIPLSKISE
LGILSWRLNDWENLKKIRGARYSYMDICDVCPEKLPNVEAKLNFEELHHT
DEEIRYLEGSGYEDVRDNDQWIRVAVKGGMLVLPAGMYHRTFLDSDNYIKAMRLF
VGEFWPTPNRPHLPAKREYVEKIINRGGTQAVEAR"

BASE COUNT 292 a 203 c 249 g 236 t

ORIGIN
Query Match 48.7% Score 424.4; DB 6; Length 980;
Best Local Similarity 83.3%; Pred. No. 1.4e-108;
Matches 495; Conservative 0; Mismatches 96; Indels 3; Gaps 1;

QY 67 ACATGGAGAACAATTTCCAGGATGTTAAGCAGGAGGTGATAGAAGCATGTTACATCGATG 126
DB 137 AGATGGAGAACAATTTCCAGGATGTTAAGCAGGAGGTGATAGAAGCATGTTACATCGATG 196
QY 127 ATAGCGAAGAGACGACGAGGCTTCTCATACCGCGCAACCCAAAGAAATTCATTCTCGTTG 186
DB 197 ACAGTGAAGAGACGACGAGGCTTCTCATATCGTGAGCCCAAGAAATTCATTCTCTCTTA 256
QY 187 ATAGCTTACAGAACTAGGAGTAATCAGCTGGCGCTAAATCCTGTATACCTGGGCAATTT 246
DB 257 GCATACTTTCAGATTAGGAATATTAAAGCTGGCGCTGAATGCTGATGACTGGGGAATG 316
QY 247 GCGAGAACTGAAGAGAAATCCGGAAGCCAGAGGTTACTTATGTTGGACATTTGTGATG 306
DB 317 ATGAGAACTCAAGAAATCCGCGAGGCCAGGGGATACCTTACATGGGATATTTGTGATG 376
QY 307 TGTGCCAGAGAGAGCTGCCAAATTTATGAAATCAAGAGTTTCTTTGAAGAAACACC 366
DB 377 TGTGCCAGAGAGAGCTGCCAAATTTATGAGGCTAAGCTGAAATTTCTTTGAAGAAACACT 436
QY 367 TGCATACCGATGAAGAAATACGCTATTGTTGAAAGGAGTGATGATCTTTGATGTGAGAG 426
DB 437 TGCATACCTGATGAAGAGATACGCTATTGTTGAGGGAAGTGGATACTTCGATGTCAAGG 496
QY 427 ACCAAATGATCAGTGGATTCGTATAGCACTGAAGAAAGAGGCGATGTTGTTCTCGCTG 486
DB 497 ACCAAATGATCAGTGGATTCGTATAGCACTGAAGAAAGGCGGATGTTGTTCTGCTG 556
QY 487 CAGGATGTACCAACCGCTTTACGTTGGACACCGACACATATATCAAGGCAATCGGACTGT 546
DB 557 CGGAGATGTATCACCGCTTCACATGGCAGTGACAACTACATCAAGGCAATCGGCTCT 616
QY 547 TTGTTGGCGATCCTGTTGGACACCGCTACAAACCGTCCCATGACCATCTTCTCTGCAAGAA 606
DB 617 TTGTTGGAGAGCTGCTGAGCGCGCTACAAACCGTCCCATGACCATCTGCGGAGTAA 676
QY 607 AGGAGTTTGGCTAAACTTCTCAAGTCAGAGGTTGAAATCAAGCAGTTGAG 660
DB 677 AGGAGTATCTGAAATAAATTTATCAA---CAGGGGTGGAACACTCAAGCTGTGGAAG 727

RESULT 4
AF068332
LOCUS AF068332 980 bp mRNA linear PLN 10-JUN-1998
DEFINITION Oryza sativa submergence induced protein 2A mRNA, complete cds.
ACCESSION AF068332
VERSION AF068332.1 GI:3201968
KEYWORDS
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
1 (bases 1 to 980)
Lorbiecke, R. and Sauter, M.
Direct Submission
TITLE Submitted (26-MAY-1998) Department of General Botany, AMPII,
JOURNAL University of Hamburg, Ohnhorststrasse 18, Hamburg 22609, Germany
FEATURES
Source 1..980
/organism="Oryza sativa"
/db_xref="dbEST:D40831"
/db_xref="taxon:4530"
139..735
/note="slp2A; similar to the Oryza sativa submergence
induced protein 2 encoded by the sequence presented in
GenBank Accession Number AF050200"

[illegible]

KEYWORDS
SOURCE
ORGANISM

FLI_CDNA.
thale cress.
Arabidopsis thaliana

REFERENCE
AUTHORS

Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. (2002) In press

JOURNAL
REFERENCE
AUTHORS

2 (bases 1 to 991)
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.

TITLE
JOURNAL
REFERENCE
AUTHORS

3 (bases 1 to 991)
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.

TITLE
JOURNAL

Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
Malibu, CA 90265, USA

COMMENT

This clone sequence is one of 5,000 Ceres full-length cDNAs made
available to TIGR and Genbank. The following quality assessment of
this set was done by comparison with known proteins: two percent of
the clones are estimated to be 5'-truncated; less than one percent
are 3'-truncated; approximately two percent represent alternative
splice variants, including unspliced introns and spliced exons; one
percent may contain premature stop codons; five percent may have
frame shifts in a coding region. A sequence is considered to be
5'-truncated if it lacks the translation initiation start (ATG). A
sequence is considered to be 3'-truncated if it lacks the
C-terminal end of the encoded protein. Please note that these cDNA
sequences are derived from the Ws or Laer ecotypes and therefore
may contain polymorphisms when compared to sequences from Col-0.
Genset carried out the library production and sequencing of the
full-length clones. Ceres, Inc. carried out the clustering of the
5' sequences, selection of clones, and sequence assembly.

FEATURES
source

1. 991
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="2738"
196..795
/codon_start=1
/product="submergence induced protein 2A"
/protein_id="AA063805.1"
/db_xref="GI:21555216"
/translation="MGEYVKDGRREVIOAWYMDSEEDORLPHHKDPKEFLSLDKLAE
LGVLSWRLADNYESDEDLKIRSRGYSYMDFCVCEPKLPNYEVKVKSPFEEHLHT
DEEIRCVAGSGYFDVRDNEAWIRVWYKGMVLVPGIYHRTVDSNDYIKAMRLF
VGEPTWPNRPHDLPARKEYIDNFVKVNEGVIDASA"

BASE COUNT 304 a 174 c 244 g 269 t

ORIGIN

Query Match 35.6%; Score 310.6; DB 8; Length 991;
Best Local Similarity 71.9%; Pred. No. 1.8e-76;
Matches 406; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

QY 68 CATGGAGAACGAAATCCAGGATGGTAAGACGGAGGTGATAGAACGATGGTACATGGATGA 127
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195 CATGGGTGAAGTGGTTAAGGATGAAGAGAAAGTGTAGTCAGGATGGTATATGGATGA 254
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 128 TAGCGAAGAGGACGAGGCTTCCTCATCACCGCAACCCAAAGAAATTCATTCCTTTGA 187
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 TAGCGAAGAGGATCAGACATTCCTCACCAAGGATCCTTAAGGAGTTTCTATCTTTTGA 314
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 188 TAAGCTTACAGAACTAGGAGTAATCAGTGGCGCCCTAAATCCTGATAACTGGGAGAAATTG 247
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 CAAACTTGCAGAGCTGGGAGTACTTAGCTGGAGACTTGTGCTGATACTATGAACCGA 374
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 248 CGAAGAACTGAAGAAATCCCGGAAGCCAGAGGTTACTCTTATGTGGACATTTGTGATGT 307

Db 375 TGAGGATTTGAAAAGATCCCGGAATCTCGTGGTTACTCTTTACATGGACTTTTGTGAGGT 434
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 308 GTGCCACAGAGATGCGCAAAATTAAGAACTAAGATCAAGAGTTTCTTTTGAAGAACCT 367
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435 ATGCCCGAAAAGCTTCCAAACTATGAAGTGAAGAGCTTTTTCGAAGAACATTT 494
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 368 GCATACCGATGAAGAAATACCGTATTGTCTTTGAAGGAGTGGATCTTTGTATGTGAGAGA 427
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 495 GCACACTGATGAAGAGATCCGTTACTGCGTTGCAGGAAGCGGTACTTTGATGTTAGGGA 554
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 428 CCAAAATGATCAGTGGATTCTGATATAGCACTCAAGAAAGGAGGCATGATTGTTCTGCCTGC 487
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 555 TCGCAATGAAGCTTGGATTAGAGTATGGGTAAGAAAGGAGGTATGATAGTCTTTACCTGC 614
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 488 AGGATGTACCACCGCTTTACGTTGGACACCGCAACCTATATCAAGCAATGCGACTGTT 547
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 615 TGGATATATCATCGCTTCACCGTAGACTCAGACAATATATCAAGCAATGCGGCTTTT 674
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 548 TGTGGCGATCCTGTGTGGACACCTACAAACCGTCCCATGACCATCTTCTCGCAAGAAA 607
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 675 CGTGGGTGAACCGGTCTGGACACCATACAATCGCCACACGACCCTCCCTGCAAGAA 734
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 608 GGAGTTTTTGGCTAAACTTCTCAAG 632
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 735 AGAGTACATCGATAATTTTCGTGAAG 759
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
AX449332 AX449332 1977 bp DNA linear PAT 03-JUL-2002
LOCUS
DEFINITION Sequence 51 from Patent WO0185946.
ACCESSION AX449332
VERSION AX449332.1 GI:21698079
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Inze,D., Boudolf,V., de Veylder,L., Acosta,J.A. and Magyar,Z.
TITLE Nucleic acid molecules encoding plant cell cycle proteins and uses
therefor
JOURNAL Patent: WO 0185946-A 51 15-NOV-2001;
CropDesign N.V. (BE)
FEATURES
source
1. 1977
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"

BASE COUNT 563 a 391 c 466 g 557 t

ORIGIN

Query Match 35.0%; Score 304.8; DB 6; Length 1977;
Best Local Similarity 73.9%; Pred. No. 7.9e-75;
Matches 387; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 99 GAGGTGATAGAACGATGCTACATGATGATAGCGAAGGAGGACGCTTCCTCATCAC 158
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1162 GAAGTGATTCAGCTTGGTATATGATGATAGTAGTAAGAGGATCAGAGACTTCCTCACCAC 1221
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 159 CGCGAACCCAAAGAAATTCATTCTCTGTTGATAAGCTTACAGAACTAGGAGTAATCAGCTGG 218
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1222 AAGGATCCTAAAGAGTTTGTATCGTTGGACAACTTCAGAGCTGGGAGTACTTAGCTGG 1281
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 219 CGCCTAAATCCTGATACTGGGAGAAATTCGAGAACCTGAAGAGAAATCCGCGAAGCCAGA 278
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1282 AGACTTGATGCTGATACTAATGAACCGATGAGGATTTTGAAGAAAGATCCGTAATCTCGT 1341
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 279 GGTACTCTTATGTGGACATTTGTGATGTGTCGCCAGAGAGCTGCCAAATTTATGAAC 338
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1342 GTTACTCTTACATGGAGCTTTTGTGAGGTATGCGCGGAAAAGCTTCACAACTATGAAGTG 1401
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

/translation="WVSSDKPREDVLOAWYMDSDSDORLPHHKEPFEVSLDOLAE
LGVLSKLDADHNDPELKKIREERGYYTMDVCEVCPEKLPNYEKIKSFEELHHT
DEIRFCAAGSGTFDVRDNEAWIRWVRKGMIIIPAGIYHRTFLDESNIKALRFF
VGEPTWTPYRNRNDHLPARQQYKDPKDVXSSHAVDATA"
BASE COUNT 270 a 164 c 205 g 250 t
ORIGIN

Query Match 32.8%; Score 285.6; DB 6; Length 889;
Best Local Similarity 68.3%; Pred. No. 2e-69;
Matches 396; Conservative 0; Mismatches 184; Indels 0; Gaps 0;
Qy 83 CCAGGATGGTAAGACGAGGCTAGAGAAGCATGTACATGGATGATAGCGAAGAGACCA 142
Db 46 CAGGATCCACGAGAGGATGCTTCAAGCTGGTACATGGATGATAGTGAAGATCA 105
Qy 143 GAGGCTTCTATCACCAGCCAGCAAGCAATTCATTCCTGTTGATAGCTTACAGAACT 202
Db 106 AAGACTCCCGCACACCAAGCAAGCAAGGAGTTGTCTGTTGACCAACTTGTGAACT 165
Qy 203 AGGAGTAATCAGCTGGCGCCCTAAATCTGTATGACTGGGAGAAATGGCAGAACCTGAAG 262
Db 166 TGGAGTCCTTAGCTGGAACATAGATGCTGATACCAATGAATGATCCAGAGCTGAAGAA 225
Qy 263 AATCCGGAAGCCAGAGGTTACTCTTATGTGGACATTTGTGTGCGCCAGAGAGCT 322
Db 226 GATTGCTGAAGAGCGTGGTTACACCTACATGGATGTTGTGAGGCTGCCCAAGAAAGTT 285
Qy 323 GCCAAATATGAACCTAAGATCAAGAGTTCTTTTGAAGAACACCTGTCATACCGATGAAGA 382
Db 286 GCCAAATATGAACAGAAATCAAAAGCTTCTTTGAAGAGCATCTTACACTGATGAGGA 345
Qy 383 ATACGCTATTGCTTGAAGGAGTGGATGATCTTTGATGTGAGAGACCAAAATGATCATGT 442
Db 346 GATCCGCTTTTGTGCTGCTGGAAGTGGCTATTTTGTATGTTTGAAGGATGCCAATGAAGCTT 405
Qy 443 GATTGCTATAGCACTGAAGAAAGAGGACATGTTGTTCTGCTGCAGGGATGTACACCG 502
Db 406 GATTGCTGTGTGGTCAAGAAAGAGGAGAAATGATCATCTTACCTGCCGGAATTTATCATCG 465
Qy 503 CTTTACGTTGGACACCGCAACTATATCAAGGCAATCGCACTGTTTGTGGCGATCCTGT 562
Db 466 CTTTACGCTAGATGAGAGCAACTACATTAAGGCTTTTCCGTTTGTGTCGAGCCAGT 525
Qy 563 TTGACACCTTACAAACCGTCCCATGACCATCTTCCCTGCAAGAAAGAGATTTTGGCTAA 622
Db 526 TTGGACTCCATACAAATCGTCCAAATGACCATCTCCCTGCAAGAGCAACAATATGTCAAGGA 585
Qy 623 ACTTCTCAAGTCACAAGGTGAATCAAGCAGTTGAAGGC 662
Db 586 TTTTGTGGAAGAGGATGTTAGCAGCCCATGCTGTTGATGCC 625

RESULT 13
AC027658/c 142737 bp DNA linear PLN 02-APR-2000
LOCUS
DEFINITION
Oriza sativa subsp. japonica BAC nbxb0006113, chromosome 10,
complete sequence.
ACCESSION
AC027658
VERSION
AC027658.1 GI:7363409
KEYWORDS
HTG.
SOURCE
Oriza sativa.
ORGANISM
Oriza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 142737)
AUTHORS
Liaca,V., Young,S., Kavchok,S., Clark,T., Choi,V.,
Nevill-Manning,C. and Messing,J.
TITLE
Rice Chromosome 10
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 142737)
AUTHORS
Liaca,V., Lou,A., Young,S., Charydczac,G., Choi,V.,
Nevill-Manning,C. and Messing,J.

TITLE Direct Submission
JOURNAL Submitted (31-MAR-2000) The Plant Genome Initiative at Rutgers -
Waksman Institute, Rutgers University, 190 Frelinghuysen Road,
Piscataway, NJ 08873
REFERENCE 3 (bases 1 to 142737)
AUTHORS
Liaca,V., Young,S., Kavchok,S., Clark,T., Charydczac,G., Choi,V.,
Nevill-Manning,C. and Messing,J.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2000) The Plant Genome Initiative at Rutgers -
Waksman Institute, Rutgers University, 190 Frelinghuysen Road,
Piscataway, NJ 08873
REMARK Chromosome 10
FEATURES
Location/Qualifiers
source l..142737
/organism="Oriza sativa"
/db_xref="taxon:4530"
/chromosome="10"
/clone="nbxb0006113"
BASE COUNT 42049 a 28618 c 30051 g 42019 t
ORIGIN
Query Match 29.3%; Score 255.6; DB 8; Length 142737;
Best Local Similarity 96.7%; Pred. No. 6e-61;
Matches 261; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 596 TCCTGCAAGAAAGAGGTTTTTGGCTAAACTTCTCAAGTCAGAAGGTGAAAATCAAGCAGT 655
Db 110173 TTCTTTCAGAAAGAGGTTTTTGGCTAAACTTCTCAAGTCAGAAGGTGAAAATCAAGCAGT 110114
Qy 656 TGAAGGCTTCTGAGGCTTTGTTGGCTCCTGCACATGCGGTTCTATATTCACACTGAATA 715
Db 110113 TGAAGGCTTCTGAGGCTTTGTTGGCTCCTGCACATGCGGTTCTATATTCACACTGAATA 110054
Qy 716 AGATGTCTATAGCAATGTAATTTAGCACAGTGGGTATGTCGCCACTCACCAACTTGA 775
Db 110053 AGATGTCTATAGCAATGTAATTTAGCACAGTGGGTATGTCGCCACTCACCAACTTGA 109994
Qy 776 AGTGAAGATTTAATGATTTTGTATTTCTTATGATCAATCGCATATAGCATTTCCG 835
Db 109993 AGTGAAGATTTAATGATTTTGTATTTCTTATGATCAATCGCATATAGCATTTCCG 109934
Qy 836 AAATGTCTTTTCAATTAACAGAGTGCATGA 865
Db 109933 AAATGTCTTTTCAATTAACAGAGTGAAGA 109904
RESULT 14
AY086650 759 bp mRNA linear PLN 25-JUN-2002
LOCUS
DEFINITION
Arabidopsis thaliana clone 26536 mRNA, complete sequence.
ACCESSION
AY086650
VERSION
AY086650.1 GI:21405360
KEYWORDS
FLI_CDNA.
SOURCE
thale cress.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 759)
AUTHORS
Haas,B.J., Volfovsky,N., Town,C.D., Troukhan,M., Alexandrov,N.,
Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
Full-length messenger RNA sequences greatly improve genome
annotation
Genome Biol. (2002) In press
REFERENCE 2 (bases 1 to 759)
AUTHORS
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
Full-length cDNA from Arabidopsis thaliana
Unpublished
REFERENCE 3 (bases 1 to 759)
AUTHORS
Brover,V., Troukhan,M., Alexandrov,N., Lu,Y.-P., Flavell,R. and
Feldmann,K.
TITLE
Direct Submission

Qy 589 ACCATCTTCTGCAAGAGGAGTTTTTGGCTAAACTTCTCAAGTCAGAGGTGAAAAATC 648
| ||| ||| ||| ||| ||| ||| | | | | | | | | | | | |
Db 499 AAGATCATCCAGCAGGAGGAGTACATCAAGAGTGTCTGAAAGAGTAGGAGTGCCTC 558
Qy 649 AAGCAG 654
| | |
Db 559 TTACAG 564

Search completed: December 1, 2002, 21:04:13
Job time : 3176 secs

	1	872	100.0	872	22	AAD15647	Rice submergence-i
	2	424.4	48.7	980	22	AAD15648	Rice submergence-i
	3	339.2	38.9	933	22	AAD15652	Cotton submergence
	4	313.8	36.0	989	21	AC484766	Arabidopsis thalia
	5	313.8	36.0	1539	21	AAC484247	Arabidopsis thalia
	6	312.4	35.8	935	21	AC484764	Arabidopsis thalia
	7	311.8	35.8	774	22	AAD15649	Tomato submergence
	8	310.6	35.6	991	21	AC434654	Arabidopsis thalia
	9	307	35.2	849	24	ABW98861	Arabidopsis thalia

PA (CROP-) CROPDESIGN NV.
XX

ALIGNMENTS

PI Sauter MM, Lorbiecke R;
 XX WPI: 2001-529915/58.
 DR P-PSDB; AAE08929.
 XX
 PT New genetic constructs, comprising the SH2A or SH2A-like genes, useful
 for altering the growth and adaptation of plants under hypoxic
 conditions, e.g. improving water logging tolerance or regulating
 anaerobic response in plants -
 XX
 PS Example 2; Page 90-91; 107pp; English.
 XX
 CC The present invention relates to genetic constructs, which comprise a
 nucleotide sequence for an SH2A or SH2A-like gene, or an SH2A or SH2A-
 CC like protein gene promoter, where the SH2A or SH2A-like gene is operably
 CC linked to a promoter sequence that directs expression of the nucleotide
 CC sequence and the SH2A or SH2A-like protein gene promoter is operably
 CC linked to a heterologous coding sequence. The SH2A gene, initially
 CC identified in rice plants, is a submergence-induced early response gene
 CC which is involved in the adaptation of plants to hypoxic conditions
 CC (such as those encountered during submersion). In addition, the SH2A gene
 CC product has also been found to belong to a family of highly conserved
 CC proteins which occur ubiquitously in eukaryotes. The genetic constructs
 CC and methods are useful for altering the growth and adaptation of plants
 CC under hypoxic conditions. In particular, these are useful for improving
 CC water logging tolerance and regulating anaerobic response in plants.
 CC These are particularly useful in horticultural, agricultural, medical,
 CC fermentation and cell culture industries. The present sequence is rice
 CC SH2A cDNA.
 XX
 SQ Sequence 872 BP; 267 A; 176 C; 206 G; 223 T; 0 other;

Query Match 100.0%; Score 872; DB 22; Length 872;
 Best Local Similarity 100.0%; Pred. No. 7.2e-260;
 Matches 872; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAGCAACAAAAACAGATCCATCCGATATCGAGGCTTCGCTTCCTCCACCCCG 60
 DB 1 AGAGCAACAAAAACAGATCCATCCGATATCGAGGCTTCGCTTCCTCCACCCCG 60

QY 61 CAATCCACATGGAGAGACGATTCAGGATGGTAAGACGAGGATAGAGCATGGTACA 120
 DB 61 CAATCCACATGGAGAGACGATTCAGGATGGTAAGACGAGGATAGAGCATGGTACA 120

QY 121 TGGATGATCGGAGAGGACGAGGCTTCCTCATCCCGGACCCCAAGAAATTCATTC 180
 DB 121 TGGATGATCGGAGAGGACGAGGCTTCCTCATCCCGGACCCCAAGAAATTCATTC 180

QY 181 CTGTTGATAAGCTTACAGAACTAGGAGTAATCAGCTGGCGCCTAAATCCTGATACTGGG 240
 DB 181 CTGTTGATAAGCTTACAGAACTAGGAGTAATCAGCTGGCGCCTAAATCCTGATACTGGG 240

QY 241 AGAATTCGAGAACCTGAGAGAAATCCGGAAGCCAGAGGTTACTCTTATGTGGACATTT 300
 DB 241 AGAATTCGAGAACCTGAGAGAAATCCGGAAGCCAGAGGTTACTCTTATGTGGACATTT 300

QY 301 GTGATGTGTCGAGAGAGCTGCCAAATATGAAACTAAGATCAAGAGCTTCTTTGAAG 360
 DB 301 GTGATGTGTCGAGAGAGCTGCCAAATATGAAACTAAGATCAAGAGCTTCTTTGAAG 360

QY 361 AACACCTGCATACCGATGAAGAAATACGCTATTGCTTGAAGGGAGTGGATACCTTTGATG 420
 DB 361 AACACCTGCATACCGATGAAGAAATACGCTATTGCTTGAAGGGAGTGGATACCTTTGATG 420

QY 421 TGACAGACCAAAATGATCAGTGTATCGTATACCATCTGAGAAAGGAGCATGATGTC 480
 DB 421 TGACAGACCAAAATGATCAGTGTATCGTATACCATCTGAGAAAGGAGCATGATGTC 480

QY 481 TGCCTGCAGGGATGTACACCGCTTTACGTTGGACACCGACCACTATATCAAGGCAATGC 540
 DB 481 TGCCTGCAGGGATGTACACCGCTTTACGTTGGACACCGACCACTATATCAAGGCAATGC 540

QY 541 GACTGTTTGTGGCGATCTCTGTTGGACACCCCTACAAACCGTCCCATGACCATCTTCCTG 600

DB 541 GACTGTTTGTGGCGATCTCTGTTGGACACCCCTACAAACCGTCCCATGACCATCTTCCTG 600
 QY 601 CAAGAAGAGATTTTGGCTAAACTTCTCAAGTCAGAAGGTGAAAATCAAGCAGTTGAAG 660
 DB 601 CAAGAAGAGATTTTGGCTAAACTTCTCAAGTCAGAAGGTGAAAATCAAGCAGTTGAAG 660
 QY 661 GCTTCTGAGGGTTTGTGGGCTCCTGCACTGCGGTCTCTATATTCACACCTGAATAAGATG 720
 DB 661 GCTTCTGAGGGTTTGTGGGCTCCTGCACTGCGGTCTCTATATTCACACCTGAATAAGATG 720

QY 721 TGTATAGCAATTAATTTAGCACAGTGGCTATGTCGCCACTCACCACCTTGAAGTGA 780
 DB 721 TGTATAGCAATTAATTTAGCACAGTGGCTATGTCGCCACTCACCACCTTGAAGTGA 780

QY 781 AAGATTTAATGATTTTGTAAATCTTATGTCATCAATCGGCATATAGCATTTCCGAAATG 840
 DB 781 AAGATTTAATGATTTTGTAAATCTTATGTCATCAATCGGCATATAGCATTTCCGAAATG 840

QY 841 TGTTCATTAACAGAGGATTCATGAAGCTGAA 872
 DB 841 TGTTCATTAACAGAGGATTCATGAAGCTGAA 872

RESULT 2
 AAD15648
 ID AAD15648 standard; cDNA; 980 BP.
 XX
 AC AAD15648;
 XX
 DT 15-NOV-2001 (first entry)
 XX
 DE Rice submergence-induced early response gene SH2B (EST S2993).
 XX
 KW Rice; SH2A; submergence induced early response gene; anaerobic response;
 KW water logging tolerance; hypoxic environment; horticultural;
 KW agricultural; medical; fermentation; cell culture industry;
 KW genetic construct; transgenic plant; EST S2993; expressed sequence tag;
 KW ss.
 XX
 OS Oryza sativa.
 XX
 FH Key Location/Qualifiers
 FT CDS 139..735
 FT /*tag= a
 FT /product= "Rice SH2B protein"
 XX
 PN WO200161020-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 16-FEB-2001; 2001WO-IB00745.
 XX
 PR 18-FEB-2000; 2000US-0183572.
 XX
 PA (CROP-) CROPDESIGN NV.
 PI Sauter MM, Lorbiecke R;
 XX
 DR WPI: 2001-529915/58.
 DR P-PSDB; AAE08930.
 XX
 CC New genetic constructs, comprising the SH2A or SH2A-like genes, useful
 for altering the growth and adaptation of plants under hypoxic
 conditions, e.g. improving water logging tolerance or regulating
 anaerobic response in plants -
 XX
 PS Example 3; Page 92-93; 107pp; English.
 XX
 CC The present invention relates to genetic constructs, which comprise a
 nucleotide sequence for an SH2A or SH2A-like gene, or an SH2A or SH2A-
 CC like protein gene promoter, where the SH2A or SH2A-like gene is operably
 CC linked to a promoter sequence that directs expression of the nucleotide

CC	sequence and the SH2A or SH2A-like protein gene promoter is operably linked to a heterologous coding sequence. The SH2A gene, initially identified in rice plants, is a submergence-induced early response gene which is involved in the adaptation of plants to hypoxic conditions (such as those encountered during submersion). In addition, the SH2A gene product has also been found to belong to a family of highly conserved proteins which occur ubiquitously in eukaryotes. The genetic constructs and methods are useful for altering the growth and adaptation of plants under hypoxic conditions. In particular, these are useful for improving water logging tolerance and regulating anaerobic response in plants. These are particularly useful in horticultural, agricultural, medical, fermentation and call culture industries. The present sequence is a rice SH2A-like cDNA, SH2B, obtained from EST (expressed sequence tag) S2993.
XX	
SQ	Sequence 980 BP; 292 A; 203 C; 249 G; 236 T; 0 other;
	Query Match 48.7%; Score 424.4; DB 22; Length 980;
	Best Local Similarity 83.3%; Pred No. 5.3e-121;
	Matches 495; Conservative 0; Mismatches 96; Indels 3; Gaps 1
QY	67 ACATGGAGACGAATTCCAGGATGGTAGAGCGGAGGTGATGAAGCATGGTACATCGGATG 126
DB	137 AGATGGAGAACCAGTTCCAGGATGCCAAGGAGGAGGTCTATCGAACGTTGGTACATCGGATG 196
QY	127 ATAGCGAAGAGGACCAGAGGCTTCCTCATCACC GGCAACCCAAAGAAATTCATCTGTTTG 186
DB	197 ACAGTGAAGAGGACCAGAGGCTTCCTCATCATCGTAGCGGACCCAAAGAAATTCATCTCTTA 256
QY	187 ATAAGCTTACAGAACTAGGAGTAATCAGCTGGCGCTTAATCTGATAACTCGGGAGAAATT 246
DB	257 GCAAACCTTCAGAGTTAGGAAATATTAACTGGCGCTGAANTGCTGATGACTGGGGAATG 316
QY	247 CGGAGAACCTCAAGAGATCCGGAAGCCAGAGGTTACTCTATCTGGACATTTGTGATG 306
DB	317 ATGAGAACCTTCAAGAAAATCCGTGAGGCGGGGATCTCTTACATGATATTTGTGATG 376
QY	307 TGTCGCCAGAGAAAGTCGCAAAATTATGAACCTAAGATCAAGAGTTTCTTTGAAGAACACC 366
DB	377 TGTGTCAGAANAAGCTGCCAAACTATGAGGCTAAGCTGAAAAATTTCTTGAAGAACACT 436
QY	367 TGCATACCGATGAAGAAATACGCTATTGCTTTGAAGGAGTGGATACATTGATGTGAGAG 426
DB	437 TGCATACTGATGAAGAGATAGCGTATTGCTTTGAGGGAAGTGGATACATTCGATGTCAAGG 496
QY	427 ACCAAAATGATCAGTGGATTGCTATAGCACATGAAGAAAGGAGGCATGATGTTCTGCCTG 486
DB	497 ACCAAAATGATCAGTGGATCCGTGTAGCAGTGAAGAAAGGGGCGCATGATGTTTGTCCCTG 556
QY	487 CAGGGATGTACACCGCTTTACGTTGGACACCGACAACCTATATCAAGGCAATGCGCACTGT 546
DB	557 CGGANAGTATCACCGCTTCACATTGGACATGACAACTACATCAAGGCAATGCGGCTCT 616
QY	547 TTGTTGGCGATCCTGTTTGGACACCCCTACAACCGTCCCCATGACCATCTTCTCTGCAAGAA 606
DB	617 TTGTGGGAGAGCCTGTCTGGAGCGCGTACAACCGTCCCCATGACCATCTGCCAGCTAGAA 676
QY	607 AGGAGTTTTTGGCTAAACTCTTCAAGTCAGAAGGTTGAAATCAAGCATTTTGAAG 660
DB	677 AGGAGTATGTCGAAAAAATATCAA---CAGGGGTGGAACTCAAGCTGTCGAG 727

Db 297 GAGGAGAGATAAAATTTCTTCAAGAACATATTCATACGTGATGAGGATCCGTAC 356
Qy 393 TGTCTTCAAGGAGCTGGATACCTTTTGATGTGAGACCAAAATGATCAGTGGATTTCGTATA 452
Db 357 TGTCTGGCAGGAAGCTGGTTATTTGATGTACGGGATCATATGATGATGATTCGTGTG 416
Qy 453 GCACCTGAAGAAAGGAGCATGATTGTTCTGCTGCTGAGGATGTACCAACCGCTTTACCTTG 512
Db 417 TGGCTGAAGAAAGGAGCATGATGTTTACCTGCTGCAATTTATCATCGCTTTACTCTG 476
Qy 513 GACACCGACACATATATCAAGGCAATCGGACTGTTTGTGCGGATCCTGTTGGACACCC 572
Db 477 GATACAGACACATATATTAAGGCAATCGGCTCTTTGCTGATCAATTTGGACTCCG 536
Qy 573 TACAACCGTCCCCATGACCATCTTCCCTGCAAGAAAGGATTTTGGCTAAACTTCTCAAG 632
Db 537 TACANTGTCGGCAGCATCATCTTCCCTGCAAGGAAGGATATATCAAGAAGCTTTTTCGG 596
Qy 633 TCAGAAGGTGAAATCAAGCAGTTGAAG 660
Db 597 GAGGAAGTGTGTCACCAAGCGTTGATG 624

RESULT 4

AAC48266

ID AAC48266 standard; DNA; 989 BP.

AC AAC48266;

XX

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 56860.

XX

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 03-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152263.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 36.0%; Score 313.8; DB 21; Length 989;
Best Local Similarity 72.2%; Pred. No. 1.le-86;
Matches 408; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 68 CATGGAGAAGCAATTCAGGATGTTAAGACGGAGGTGATAGAGCATGTACATGGATGA 127
 ||||| ||||| ||||| ||||| |||||
Db 193 CATGGTGAAGTGGTTAAGATGGAAGAAGATGATCCAGGCATGTTATATGATGA 252
 ||||| ||||| ||||| ||||| |||||
QY 128 TAGCGAAGAGGACAGAGGCTTCCTCATCCCGCGCAACCCAAAGAAATTCATTCCTGTGA 187
 ||||| ||||| ||||| ||||| |||||
Db 253 TAGCGAAGAGGATCAGAGACTTCCTCACCACAAGACTCTAAGGAGCTTCTATCTTTGGA 312
 ||||| ||||| ||||| ||||| |||||
QY 188 TAAGCTTACAGAACTAGGAGTAAATCAGCTGGCGCCTAAATCTCTGATGTAATCTGGGAATTC 247
 ||||| ||||| ||||| ||||| |||||
Db 313 CAAACTTGCAGAGCTGGGAGTACTTAGCTGGAGACTTGATGCTGATAACTATGAAACCGA 372
 ||||| ||||| ||||| ||||| |||||
QY 248 CGAGAACCTGAAGAGAAATCCGCGAAGCCAGAGGTTACTCTTATGTGGACATTTGTGATGT 307
 ||||| ||||| ||||| ||||| |||||
Db 373 TGAGGATTTGAAAAGATCCGCAATCTCGTGTACTCTTACATGGACTTTTGTGAGGT 432
 ||||| ||||| ||||| ||||| |||||
QY 308 GTGCCAGAGAAGCTGCCAAATTTATGAAACTAAGATCAAGAGTTTCTTTGAACAACACCT 367
 ||||| ||||| ||||| ||||| |||||
Db 433 ATGCCCGGAAAAGCTTCCAACATATGAAGTGAAGAGCTTTTTCGAGAACAATTT 492
 ||||| ||||| ||||| ||||| |||||
QY 368 GCATACCGATGAAGAAATACGCTATTGTCTTGAAGGGAGTGGATACTTTGATGTGAGAGA 477
 ||||| ||||| ||||| ||||| |||||
Db 493 GCACACTGATGAAGATCCGTTACTGCGTTCGAGGAGCGTTACTTTGATGTGAGGA 552
 ||||| ||||| ||||| ||||| |||||
QY 428 CCAAAATGATGAGTGGATTCGTATAGCACTGAAGAAAGGAGGCATGATTTCTGCCTGC 487
 ||||| ||||| ||||| ||||| |||||
Db 553 TCGCAATGAAGCTTGGATTAGATGTTGGGTAAAGAAAGGAGGTATGATAGTCTTACCTGC 612
 ||||| ||||| ||||| ||||| |||||
QY 488 AGGATGTACCACCGCTTTACGTTGGACACCGCACTATATCAAGGCATCGACTGTT 547
 ||||| ||||| ||||| ||||| |||||
Db 613 TGAATATATATCATCGCTTCACCGTAGACTCAGACAATATATCAAGGCAATCGCGCTTT 672
 ||||| ||||| ||||| ||||| |||||
QY 548 TGTGTGGCATCTGTTTGGACACCTTACAAACCGTCCCATGACCATCTTCTCTGCAAGAAA 607
 ||||| ||||| ||||| ||||| |||||
Db 673 CGTTGTGAACCGCTCTGGACACCATACAAATCCCCACAGCACCACTCCCTGCAAGGAA 732
 ||||| ||||| ||||| ||||| |||||

QY 608 GGAGTTTTTGGTAAACTTCTCAAG 632
 ||||| ||||| |||||
Db 733 AGAGTACATCGATAATTTCTGTGAAG 757

RESULT 5
AAC48247
ID AAC48247 standard; DNA; 1539 BP.
XX AAC48247;
AC AAC48247;
XX
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 56793.
DE
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX

PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 03-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

[illegible]

Db 558 TGGCGCTGCTGCTGCTGTTAATG 580
RESULT 8
AAC43654
ID AAC43654 standard; DNA; 991 BP.
XX AC AAC43654;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 40010.
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway;
XX KW metabolic pathway; promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.

PR	18-AUG-1999;	99US-0149426.	
PR	20-AUG-1999;	99US-0149722.	
PR	20-AUG-1999;	99US-0149723.	
PR	20-AUG-1999;	99US-0149929.	
PR	23-AUG-1999;	99US-0149902.	
PR	23-AUG-1999;	99US-0149930.	
PR	25-AUG-1999;	99US-0150566.	
PR	26-AUG-1999;	99US-0150884.	
PR	27-AUG-1999;	99US-0151065.	
PR	27-AUG-1999;	99US-0151066.	
PR	27-AUG-1999;	99US-0151080.	
PR	30-AUG-1999;	99US-0151303.	
PR	31-AUG-1999;	99US-0151438.	
PR	01-SEP-1999;	99US-0151930.	
PR	07-SEP-1999;	99US-0152363.	
PR	10-SEP-1999;	99US-0153070.	
PR	13-SEP-1999;	99US-0153758.	
PR	15-SEP-1999;	99US-0154018.	
PR	16-SEP-1999;	99US-0154039.	
PR	20-SEP-1999;	99US-0154779.	
PR	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match			35.6%;
Best Local Similarity			71.9%;
Matches 406; Conservative			0;
Mismatches 159; Indels			0;
Gaps			0;
QY	68	CATGGAGAACGAATTCACGAGTGTAGACGGAGGTGATAGAACGATGGTACATGGATGA	127
Db	195	CATGGGTGAAGTGGTTAAGGATGGAAGAGAAAGTATCCAGGCATGTATATGGATGA	254
QY	128	TAGCGAGAGGACCGAGGCTTCCTCATCCACCGCAACCCAAAGAAATTCATTCCTGTTGA	187
Db	255	TAGCGAAGAGGATGAGAGACTTCCTCACCACAAAGGATCCTTAAGAGGATTCCTATCTTTTGA	314
QY	188	TAAGCTTACAGAACTAGGAGTAATCAGCTGGCGCCTAAATCCTGATAACTCGGAGAATTG	247

Db	315	CAAACTTGCAGAGCTGGAGTACTTAGCTGGAGACTTTGATGCTGATAACTATGAAACCGA	374
QY	248	CGAGAACCTCAAGAGATCCGCGAAGCCAGAGGTACTCTTATCTGGACATTTGTGATGT	307
Db	375	TGAGGATTTGAAAAAGATCCGGAATCTCGTGGTTACTCTTACATGGACTTTTGTGAGGT	434
QY	308	GTGCCAGAGAAAGCTGCGCAAAATTAAGATCAAGAGTTTCTTTTGAAGAACAACCT	367
Db	435	ATGCCCGGAAAAGCTTCCAAACTATGAAGTGAAGTGAAGAGCTTTTTCGAAGAACATTT	494
QY	368	GCATACCGATGAAGAAATACGCTATTGCTTTGAAGGAGTGGATACATTTTGTATGTGAGAGA	427
Db	495	GCACACTGATGAAGAGATCGGTTACTCGTTTGCAGGAAGCGGTACTTTGTATGTAGGGA	554
QY	428	CCAAATGATCAGTGGATTCGTATAGCACTGAAGAAAGGAGGATGATTCTTCTGCCTGC	487
Db	555	TCCCAATGAAGCTTGGATTAGAGTATGGTAAAGAGGAGGTATGATAGTCTTACCTGC	614
QY	488	AGGATGTACCAACCGCTTACGTTTGGACACCGCAACTATATCAAGCAATGGCACTGTT	547
Db	615	TGGAATATATCATCGCTTCACCGTAGACTCAGACAACTATATCAAGCAATGGCGCTTT	674
QY	548	TGTTGGCGATCCTTTTGGACACCTACACCGTCCCGCATGACCATCTTCTTGCAGAAA	607
Db	675	CGTGGTGAACCGGTCTGGACACCATACAATCGCCACACGACCATCTCCTTGCAGGAA	734
QY	608	GGAGTTTTTGGCTAAACTTCTCAAG	632
Db	735	AGATACATCGATAATTCGTGAAG	759
RESULT 9			
ABN98861/G			
ID	ABN98861 standard; DNA; 849 BP.		
XX	ABN98861;		
XX	01-AUG-2002 (first entry)		
DE	Arabidopsis thaliana expressed polynucleotide SEQ ID NO 629.		
KW	Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;		
KW	disease; crop; thale cress; tolerance factor; insect; pathogen;		
XX	Arabidopsis thaliana.		
OS	US2002023281-A1.		
XX	21-FEB-2002.		
PF	26-JAN-2001; 2001US-0770445.		
XX	27-JAN-2000; 2000US-178472P.		
PA	(GORL/) GORLACH J.		
PA	(ANY/) AN Y.		
PA	(HAM/) HAMILTON C M.		
PA	(PRIC/) PRICE J L.		
PA	(RAIN/) RAINES T M.		
PA	(YUY/) YU Y.		
PA	(RAME/) RAMEAKA J G.		
PA	(PAGE/) PAGE A.		
PA	(MATH/) MATHIEW A V.		
PA	(LEDF/) LEDFORD B L.		
PA	(WOES/) WOESSNER J P.		
PA	(HAAS/) HAAS W D.		
PA	(GARC/) GARCIA C A.		
PA	(KRIC/) KRICKER M.		
PA	(SLAT/) SLATER T.		
PA	(DAVI/) DAVIS K R.		
PA	(ALLE/) ALLEN K.		

(HOFF/) HOFFMAN N.
(HURB/) HURBAN P.

Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
Garcia CA, Krickler M, Slater T, Davis KR, Allen K, Hoffman N;
Hurban P;

WPI; 2002-400781/43.

New Arabidopsis thaliana nucleic acid for identifying homologous genes,
producing compositions that modulate the expression or function of its
encoded protein, and mapping functional regions of protein -

Claim 1; SEQ ID NO 629; 49pp + Sequence Listing; English.

The invention relates to an Arabidopsis thaliana nucleic acid (I)
comprising a sequence capable of hybridising under stringent conditions
to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),
given in the specification or its fragment. A polypeptide (II) encoded by
(I), a transgenic plant (III) comprising an exogenous nucleic acid or a
genetically modified cell (IV) comprising an exogenous nucleic acid, is
useful for screening a candidate agent for its biological effect. (I) is
useful in identifying homologous or related genes, in producing
compositions that modulate the expression or function of its encoded
protein, mapping functional regions of the protein and in studying
associated physiological pathways. (I) is also useful for the genetic
manipulation of cells, particularly plant cells. (I) is also useful in
screening assays of various plant strains to determine the strains that
are best capable of withstanding a particular disease or environmental
stress. (II) and (III) are useful for screening of biologically active
agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
pathways. The screened agents are useful in improved methods of treating
crops to prevent or treat disease. (II) are also useful in screening
programs to identify agents that mimic or enhance the action of tolerance
factors. Such agents are useful in improved methods of treating crops to
enhance their tolerance to environmental stress. (I) is also useful
for enhancing or inhibiting production of a biosynthetic product in a
plant. (III) is useful for identifying other mediators that may induce
expression of proteins of interest, for establishing the extent to which
any specific insect and/or pathogen is responsible for damage to a
particular plant, for identifying other mediators that enhance or induce
tolerance to environmental stress, for identifying factors involved in
biosynthetic pathways of nutritional, commercial, or medicinal value and
for identifying productions of nutritional, commercial or medicinal
value. (IV) is useful in the study of genetic function and regulation,
for alteration of the cellular metabolism and for screening compounds
that may affect the biological function of the gene or gene products.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from USPTO
at seqdata.uspto.gov/sequence.html?DocID=999909770445.

Sequence 849 BP; 252 A; 190 C; 151 G; 247 T; 9 other;

Query Match 35.2%; Score 307; DB 24; Length 849;
Best Local Similarity 72.4%; Pred. No. 1.3e-84;
Matches 391; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY 83 CCAGGATGTAGACGGAGGTGATAGACGATGGTATACATGGATGATAGCGAAGACCA 142
DB 849 CAAGGATGGAGAGAAGTATTCAAGCTGGTATATGGATGATAGTGAAGAGATCA 790
QY 143 GAGGCTTCATCCCGCGACCCNAAAGTAATTCATCTCTGTGATAAGCTTACAACT 202
DB 789 GAGACTTCCTCACCAAGAGATCCTAAAGAGTTGTATCGTTGGCAAACTTGCAGAGT 730
QY 203 AGAGTAATACGTCGCGCTAAATCTCTATAACTGGGAGAATTGGCAAGACTGAAGAG 262
DB 729 GGGAGTACTAGCTGGAGACTTGTATGCTGATTAACACCGGATGAGGATTTGAATA 670
QY 263 AATCCCGAAGCGAGGTTACTTCTTATGTGGACATTTGTGTGTCGCCAGAGAAGCT 322
DB 669 GATCCGTGATCTCGGGTACTCTTACATGGACTTTTGTGAGGTATGCCCGAAGACT 610

QY 323 GCCAAATTATGAACTAAGATCAAGAGTTTCTTTGAGAAGACACCTGCATACCGATGAAGA 382
DB 609 TCCAAACTATGAAGTGAAGTAAAGAGCTTTTTCGAAGAACAATTTGCATCTACTGATGAAGA 550
QY 383 AATACGCTATTGCTTTGAAGGAGTGGATACCTTTGATGTGAGAGACCAAAATGATCAGTG 442
DB 549 GATCCGTTACTCGCTTGCAGGAAGTGGTTACTTTGATGTGAGAGATGCAATGAAGCTTG 490
QY 443 GATTCGTATACACTGAAGAAAGGAGGATGATTCTTCCTGCCTGCAGGATGTACCACG 502
DB 489 GATTAGGTATGGTAAANNNNNNNAGCTATGATANNNTTACCTGCTGGGATCTATCATCG 430
QY 503 CTTTACGTTGGACACCGACAACCTATATCAAGCAATGGGACTGTTTGTGGCGATCCTGT 562
DB 429 CTTACCGTGGACTCTGCAACTATATCAAGCAATGCGGCTTTTCGTGGGTGAACCGT 370
QY 563 TTGGACACCTACAAACCGTCCCATGACCACTCTCTGCAAGAAAGGAGTTTGTGGCTAA 622
DB 369 ATGGACACCATACAATCGCCACACGACCATCTCTCTGCAAGAAAGAAATATGTCGATAA 310
RESULT 10
ID AAS96322 standard; cDNA; 1977 BP.
XX AAS96322;
AC AAS96322;
XX 26-FEB-2002 (first entry)
DT 26-FEB-2002 (first entry)
XX Arabidopsis cDNA encoding cell cycle protein CCP22.
DE Arabidopsis cDNA encoding cell cycle protein CCP22.
XX Cell cycle protein; CCP; ss; cell cycle regulation; herbicide;
KW plant growth regulator; plant development; abiotic stress; biotic stress;
KW nutrient deprivation; pathogen attack; crop yield.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX WO200185946-A2.
PN 15-NOV-2001.
PD 14-MAY-2001; 2001WO-IB01307.
PF 12-MAY-2000; 2000US-204045P.
PR (CROP-) CROPDESIGN NV.
XX Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;
PI WPI; 2002-062249/08.
XX P-PSDB; AAU72532.
DR New cell cycle protein and nucleic acid molecule encoding it useful for
PT regulating cell cycle progression in plants and for identifying
PT modulators which are useful as herbicides or plant growth regulators -
XX Claim 38; Fig 22; 316pp; English.
PS The invention relates to a novel cell cycle protein (CCP) and the
XX polynucleotides encoding them. CCP is useful for identifying a compound
CC which modulates the activity of the polypeptide and which binds to the
CC polypeptide and an anti-CCP antibody is useful for detecting the presence
CC of CCP in a sample. A CCP modulator is useful for modulating the cell
CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,
CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola.
CC CCP nucleic acid and polypeptide molecules are useful as modulating
CC agents in regulating cell cycle progression in plants. CCP is useful to
CC treat disorders characterised by insufficient or excessive production of
CC CCP protein or production of CCP protein forms which have decreased or
CC aberrant activity. Compounds that bind to or modulate the activity
CC of CCP polypeptide are useful as herbicides or plant growth regulators.
CC The polynucleotide is useful for modifying cell fate, plant development,

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132488.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-016192.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	

Query Match 33.1%; Score 288.2; DB 21; Length 597;
Best Local Similarity 70.8%; Pred. No. 7.2e-79;
Matches 405; Conservative 0; Mismatches 149; Indels 18; Gaps 1;

Qy	69	ATGGAAGCAATTCAGGATGGTAAGACGGAGGTGATAGAAAGCATGTACATGGATGAT 128
Db	1	ATGGGTGAAGCGGTCAAGGATGGGANAGAAAGTGATTCAAGCTTGGTATATGGATGAT 60
Qy	129	AGCGAAGAGGACGAGAGGCTTCCTCATCACC CGCAACCCAAAGAAATTCATTCCTGTTGAT 188
Db	61	AGTGAAGAGGATCAGACACTTCCTCACCACAAGGATCCTAAAGAGTTTGTATCGTTGGGC 120
Qy	189	AAGCTTACAGAACTAGGAGTAATCAGCTGGCGCCTAAATCCTGATAACTGGGAGAAATGCC 248
Db	121	AAACTTGCAGAGCTGGGAGTACTTAGCTTGGAGACTTGTATGCTGATTAACACGAAACCGAT 180
Qy	249	GAGAACTGAAGAGAAATCCGCGAAGCCAGAGGTTACTCTTA-----T 290
Db	181	GAGGATTTGAAAAAGATCCGTGAATCTCGTGGTTACTCTTACATGAAAAATGTTCAAAT 240
Qy	291	GTGGACATTTGTATGTGTGCCAGAGAGCTGCCAAATTTATGAACATAGATCAAGAGT 350
Db	241	CAGNACTTTTGTGAGGTATGCCGCGAAAGCTTCCAAACTATGAAGTGAAGCTAAAGAGC 300
Qy	351	TTCTTTGAAGAACACCTGCATACCGATGAAGAAATACGCTATTGTCTTGAAGGGAGTGA 410
Db	301	TTTTTCGAAGAACATTTGCATACTGATGAAGAGATCCGTTACTGCGTTGCAGGAAGTGT 360
Qy	411	TACTTTGATGTGAGAGACCAAAATGATCAGTGGATTTCGTATAGCACTGAAGAAAGAGGC 470
Db	361	TACTTTGATGTGAGAGATCGCAATGAAGCTTGGATTAGGTTAGGTTAAAGAGGAGCT 420
Qy	471	ATGATTTGTTCCCTGAGGAGGTGTACCACCGCTTTAGCTTGGACACCGCAACTATATC 530
Db	421	ATGATAGTCTTACCTGTGGGATCTATCATCGCTTTCACCGTGGACTCTGCACACTATATC 480
Qy	531	AAGCAATGCGAGTCTTGTGTGGCGATCCTGTTTGGACACCCCTACAACCGTCCCCATGAC 590
Db	481	AAGGCAATGCGGCTTTTCGTGGGTGAACCGGTATGGACACCATACAAATCGCCACACGAC 540
Qy	591	CATCTTCTCGAAGAAAGAGTTTTTTGGCTAA 622

Db	541	CATCTTCTCGAAGAAAGATATGTCGATAA 572
RESULT 14		
AD	AD15651	
ID	AD15651	standard; DNA; 889 BP.
XX		
AC	AD15651;	
XX		
DT	15-NOV-2001	(first entry)
XX		
DE	Soybean submergence-induced early response gene SH2A-like DNA.	
XX		
KW	Soybean; SH2A-like protein; submergence induced early response gene;	
KW	anaerobic response; water logging tolerance; hypoxic environment;	
KW	horticultural; agricultural; medical; fermentation;	
KW	cell culture industry; genetic construct; transgenic plant; ds.	
XX		
OS	Glycine max.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	2..634
FT		/*tag= a
FT		/product= "Soybean SH2A-like protein"
FT		/note= "CDS does not include start codon"
FT		/partial
XX		
PN	WO200161020-A2.	
XX		
PD	23-AUG-2001.	
XX		
PF	16-FEB-2001; 2001WO-IB00745.	
XX		
PR	18-FEB-2000; 2000US-0183572.	
XX		
PA	(CROP-) CROPDESIGN NV.	
XX		
PI	Sauter MM, Lorbiecke R;	
XX		
DR	WPI; 2001-529915/58.	
DR	P-PSDB; AAE08933.	
XX		
PT	New genetic constructs, comprising the SH2A or SH2A-like genes, useful	
PT	for altering the growth and adaptation of plants under hypoxic	
PT	conditions, e.g. improving water logging tolerance or regulating	
PT	anaerobic response in plants -	
XX		
PS	Claim 31; Page 98-99; 107pp; English.	
XX		
CC	The present invention relates to genetic constructs, which comprise a	
CC	nucleotide sequence for an SH2A or SH2A-like gene, or an SH2A or SH2A-	
CC	like protein gene promoter, where the SH2A or SH2A-like gene is operably	
CC	linked to a promoter sequence that directs expression of the nucleotide	
CC	sequence and the SH2A or SH2A-like protein gene promoter is operably	
CC	linked to a heterologous coding sequence. The SH2A gene, initially	
CC	identified in rice plants, is a submergence-induced early response gene	
CC	which is involved in the adaptation of plants to hypoxic conditions	
CC	(such as those encountered during submersion). In addition, the SH2A gene	
CC	product has also been found to belong to a family of highly conserved	
CC	proteins which occur ubiquitously in eukaryotes. The genetic constructs	
CC	and methods are useful for altering the growth and adaptation of plants	
CC	under hypoxic conditions. In particular, these are useful for improving	
CC	water logging tolerance and regulating anaerobic response in plants.	
CC	These are particularly useful in horticultural, agricultural, medical,	
CC	fermentation and cell culture industries. The present sequence is a	
CC	soybean SH2A-like DNA.	
XX		
SQ	Sequence 889 BP; 270 A; 164 C; 205 G; 250 T; 0 other;	
Query Match 32.8%; Score 285.6; DB 22; Length 889;		
Best Local Similarity 68.3%; Pred. No. 5.6e-78;		
Matches 396; Conservative 0; Mismatches 184; Indels 0; Gaps 0;		

Qy	83	CCAGGATGGTAAAGACGGAGGTGATAGAGCATGGTACATGGATGATAGCGAAGGACCA	142
Db	46	CAAGCATCCACGAGAGGTGTCCTTCAAGCTGGTACATGGATGATAGAGATCA	105
Qy	143	GAGGCTTCCTCATCACCGCAACCCAAAGAATTTCATTCCTGTTGATAGCTTACAGAACT	202
Db	106	AAGACTCCCCCACCACCAAGAACCAAGAGGTTTGTCTCGTTGGACCAACTTGTGTAAC	165
Qy	203	AGGAGTAATCAGCTGGCGCCTAAATCTCTGATACTGGGAGAAATTCGCGAGAACCTGAAGAG	262
Db	166	TGAGTCCCTTAGCTGGAAACTAGATGCTGATAACCAATGATCAAGAGCTGAAGAA	225
Qy	263	AATCCGCGAAGCCAGAGGTTTACTCTTATGTGGACATTTGTGCTGCGCCAGAGAAGCT	322
Db	226	GATTCGTGAAGAGCGTGGTTACACCTACATGGATGTTGTGAGGTCGCGCCAGAAAAAGTT	285
Qy	323	GCCAAATATGAACACTAAGATCAAGAGTTTCTTTGAAGAACACCTGCATACCGATGAAGA	382
Db	286	GCCAAATATGAACAGAAAAATCAAAAGCTTCTTTGAAGAGCATCTTCACACTGATGAGGA	345
Qy	383	AATACGCTATGCTTCTGAAGGAGTGGATACTTTGATGTGAGAGACCAAAATGATCAGTG	442
Db	346	GATCCGCTTTTGTGCTGCTGGAAGTGGCTATTTTGATGTTAGGATCGCAATGAAGCTTG	405
Qy	443	GATTCTATAGCACTGAAGAAAGGAGGCATGATTGTTCTGCTGCAGGGATGATACCAACG	502
Db	406	GATTCGTGTGTGGTCAAGAAAGAGGAAGTATCATCTTACCTGCGCGGAATTTATCATCG	465
Qy	503	CTTTACGTTGGACACCGACCACTATATCAAGGCAATCGACTGTTCTGCGGATCCTGT	562
Db	466	CTTTACGCTAGATGAGAGCAACTACATTAAGGCTTTCGCTTTTGTGATGAGCCAGT	525
Qy	563	TTGGACACCTACAAACCGTCCCATGACCATCTTCTGCAAGAAAGGAGTTTGTGGCTAA	622
Db	526	TTGGACTCCATACAATCGTCCAAATGACCATCTCCCTGCRAGACAAATATGTCAAGGA	585
Qy	623	ACTTCTCAAGTCAAGAGGTGAAATCAAGCAGTTGAAGGC	662
Db	586	TTTGTGGAAAGGATGTTAGCAGCCATGCTGTTGATGCC	625
RESULT 15			
AAC48250			
ID	AAC48250 standard; DNA; 1300 BP.		
XX	AAC48250;		
XX	AAC48250;		
DT	18-OCT-2000 (first entry)		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 56805.		
XX	Arabidopsis thaliana.		
XX	OS		
XX	EP1033405-A2.		
XX	06-SEP-2000.		
PD	25-FEB-2000; 2000EP-0301439.		
PPF	XX		
XX	XX		
PR	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999; 99US-0123180.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	25-MAR-1999; 99US-0126264.		
PR	29-MAR-1999; 99US-0126785.		
PR	01-APR-1999; 99US-0127462.		
PR	06-APR-1999; 99US-0128234.		
PR	08-APR-1999; 99US-0128714.		

PR	16-APR-1999;	99US-0129845.
PR	19-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	23-APR-1999;	99US-0130891.
PR	28-APR-1999;	99US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	30-APR-1999;	99US-0132407.
PR	04-MAY-1999;	99US-0132484.
PR	05-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	06-MAY-1999;	99US-0132487.
PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	02-JUL-1999;	99US-0142154.
PR	05-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144086.
PR	16-JUL-1999;	99US-0144085.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 20:05:08 ; Search time 72 Seconds
(without alignments)
3714.196 Million cell updates/sec

Title: US-09-785-738A-1
Perfect score: 872
Sequence: 1 agcgaacacacacagaat.....acaggagtcagtgaa 872

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	199.8	22.9	1107	2	US-08-933-750C-77
2	199.8	22.9	1107	3	US-09-234-613-77
3	172	19.7	1093	4	US-09-348-265-1
4	55.6	6.4	7218	1	US-08-232-463-14
5	33.4	3.8	1323	1	US-08-858-207A-50
6	32.8	3.8	1627	1	US-08-615-170-2
7	32.8	3.8	1666	1	US-08-615-170-4
8	32.2	3.7	7432	1	US-07-852-260-1
9	32.2	3.7	7432	2	US-08-461-503-1
10	32.2	3.7	7432	3	US-08-465-250-1
11	31.8	3.6	3572	2	US-08-713-815A-2
12	31.4	3.6	4325	1	US-08-453-924-2
13	31.4	3.6	4325	1	US-08-471-791-29
14	31.4	3.6	4325	5	PCT-US91-01746-29
15	31.2	3.6	2169	4	US-09-434-408-3
16	30.8	3.5	686	3	US-09-142-514-2
17	30.8	3.5	19307	3	US-08-836-032A-10
18	30.8	3.5	19307	4	US-09-427-048A-10
19	30.4	3.5	701	2	US-08-829-876-102
20	30.4	3.5	701	4	US-09-234-874A-102
21	30.4	3.5	704	2	US-08-829-876-100
22	30.4	3.5	704	2	US-08-829-876-104
23	30.4	3.5	704	4	US-09-234-874A-100
24	30.4	3.5	704	4	US-09-234-874A-104
25	30.4	3.5	3789	4	US-09-091-889A-3
26	30.4	3.5	8450	4	US-09-091-889A-5
27	30.2	3.5	574	4	US-09-221-017B-413

28	30.2	3.5	1245	4	US-08-961-527-257	Sequence 257, Appl
29	30.2	3.5	2882	1	US-08-424-567-1	Sequence 1, Appli
30	30.2	3.5	2882	2	US-08-711-928-1	Sequence 1, Appli
31	30.2	3.5	2882	4	US-09-184-937-1	Sequence 1, Appli
32	30.2	3.5	7163	4	US-08-961-527-67	Sequence 67, Appl
33	30	3.4	289	4	US-09-007-005-17	Sequence 17, Appl
34	30	3.4	289	4	US-09-244-796-17	Sequence 17, Appl
35	30	3.4	903	4	US-09-134-001C-1620	Sequence 1620, Ap
36	29.8	3.4	2520	4	US-08-961-527-14	Sequence 14, Appl
37	29.8	3.4	2520	4	US-08-936-165A-214	Sequence 214, App
38	29.8	3.4	2560	4	US-08-961-527-114	Sequence 114, App
39	29.8	3.4	4327	4	US-08-961-527-117	Sequence 117, App
40	29.8	3.4	7431	4	US-09-306-998-2	Sequence 2, Appli
41	29.8	3.4	9048	4	US-08-961-527-159	Sequence 159, App
42	29.6	3.4	248	4	US-09-007-005-32	Sequence 32, Appl
43	29.6	3.4	248	4	US-09-244-796-32	Sequence 32, Appl
44	29.6	3.4	277	4	US-09-007-005-3	Sequence 3, Appli
45	29.6	3.4	277	4	US-09-244-796-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-933-750C-77
: Sequence 77, Application US/08933750C
: Patent No. 5932442
: GENERAL INFORMATION:
: APPLICANT: Lal, Preeti
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Bandman, Olga
: APPLICANT: Shah, Purvi
: APPLICANT: Au-Young, Janice
: APPLICANT: Yue, Henry
: APPLICANT: Guegler, Karl J.
: APPLICANT: Corley, Neil C.
: TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
: NUMBER OF SEQUENCES: 98
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/933,750C
: FILING DATE: September 23, 1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0356 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-855-0555
: TELEFAX: 415-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 77:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1107 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: LINDNOT03


```

; APPLICANT: SAHARA, Hiroeki
; APPLICANT: SUZUKI, Manabu
; APPLICANT: HAMURO, Junji
; TITLE OF INVENTION: Human Gastric Cancer Antigen Gene and Gastric
; TITLE OF INVENTION: Cancer Antigen Protein
; FILE REFERENCE: OP871
; CURRENT APPLICATION NUMBER: US/09/348,265
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 10-197852
; EARLIER FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1093
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)..(534)
US-09-348-265-1

```

Query Match	19.7%	Score 172;	DB 4;	Length 1093;
Best Local Similarity	61.9%	Pred. No. 5.8e-44;		
Matches 306;	Conservative 0;	Mismatches 185;	Indels 3;	Gaps 2;
QY 122	CGATGATCGGAAGAGGACACAGAGGCTTCCTCATCCGCCGGAACCCAAAGAAATTCATTCC	181		
Db 67	GGACGACGCCCGGGCGACCGCGGCAACCCACCGCCCGCCAGCCCGGCCCGCCAGTGG	126		
QY 182	TG--TTGATAAGCTTACAGAATAGGAGTAATCAGCTGGCGCTAAATCCTGATAACTGG	239		
Db 127	GCCTTGGAGCAGCTCGCGGCTCTACTTGGNAAGTGGATGCTGACAAATAT	186		
QY 240	GAGAATTCGGAGAACCTGAAGAGAAATCCGCGAAGCCAGAGGTTACTTCTATGTGGACATT	299		
Db 187	GAGAATGTCACAGAATTAGAAAGATCCGAAGAGAGAGGAACTACTCTGGATGGACATC	246		
QY 300	TGTGATGTCGCCAGAGAACTCCCAAAATTGAAACTAAGATCAAGAGTTCCTTTGAA	359		
Db 247	ATAACCATATCGAAGATATAACTTCCAAATATTAGAAAGAAAGATTAGATGTCTACGAG	306		
QY 360	GAACACCTGCATACCCGATGAAGAATACGCTATTGTCTTGAAGGGATCGATACTTTGAT	419		
Db 307	GAGCATTTGCATTTGGACGATGAGATCCGCTACATCCTGGATGGCASTCGGTACTTCAC	366		
QY 420	GTGAGAGACCAAAATGATCATGTGGATTCGTATAGCACTGAAGAAGGAGGCATGATGTT	479		
Db 367	GTGAGGACAAGGAGGACCACTGTGATCCGGATCTTCATGTGAGAAGGAGACATGGTGACG	426		
QY 480	CTGCCTGCAGGGATGTACCCCGCTTTACGTTGGACACCCACAACTATATCAAGGCAATG	539		
Db 427	CTCCCCGGGGGATCTATCAACCGTTTCCCGTGTGACGAGAAGAACTACACGAAGGCCATG	486		
QY 540	CGAC-TGTTTGTGGCGATCTCTGTTTGGACACCCCTACAACCGTCCCATGACATCTTCC	598		
Db 487	CGSGTGTGTTGGGAGAACCGGTCTGACACGCTGACACCGGCTGACCATTTGA	546		
QY 599	TGCAAGAAGGAGT	612		
Db 547	AGCCCGCGGCAGT	560		

RESULT 4
US-08-232-463-14/c

0802324-177 Application US/08232463
? Sequence 14, Patent No. 5670367
? PATENT INFORMATION:
? APPLICANT: DORNER, F.
? APPLICANT: SCHEIFELINGER, F.
? APPLICANT: FALKNER, F. G.
? TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
? NUMBER OF SEQUENCES: 52
? CORRESPONDENCE ADDRESS:
?

ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232.463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935.313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Watch	6.4%;	Score 55.6;	DB 1;	Length 7218;
Best Local Similarity	8.4%;	Pred. No.3e-07;		
Matches	37;	Conservative	217;	Mismatches 186; Indels 0; Gaps
QY	55	ACCCGCCAATCACATGGAAGAACGAATTCAGAGTGGTAAGACGAGGAGGTGATAGAAGCAT	114	
DB	1477	ACCTATCTATGCAAGTAGTTAAAGAGATAGAAGATTGCTACRRRRRRRRRRRRRRR	1418	
QY	115	GGTACATGATGATAGCGAAGAGGACGAGCTTCCTCATCACCGCGAACCACCAAGAAAT	174	
DB	1417	RR	1958	
QY	175	TCATTCTGTGTATAAGCTTACAGAACTAGGAGTAATCAGCTGGCGCCTAAATCCTGATA	234	
DB	1357	RRR	1298	
QY	235	ACTGGGAGAAATTGCGAAGACCTGAGAGAAATCCGCGAAGCAGCAGGTTACTTCTATGTGG	294	
DB	1297	RRR	1238	
QY	295	ACATTGTGATGTGCCACAGAGAAGCTGCCAAATATCAAACTAAGATCAAGAGTTTCT	354	
DB	1237	RRR	1178	
QY	355	TTGAAGACACCTGCATACCGATGAAGAAATACGCTATTGCTTGAAGGAGTGCATGACT	414	
DB	1177	RRR	1118	
QY	415	TTGATGTGAGACCAAAATGATCATGTGGATCGTATAGCACTGAAGAAAGAGCGCATGA	474	
DB	1117	RRR	1058	
QY	475	TTGTTCTGCCTGCAGGGATG	494	
DB	1057	TCCTCGACCTGCAGCCAAG	1038	


```

RESULT 8
US-07-852-260-1
; Sequence 1, Application US/07852260
; Patent No. 5525715
; GENERAL INFORMATION:
; APPLICANT: Racanietto, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/852.260
; FILING DATE: 19920619
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7432 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 743..7361
; US-07-852-260-1

```

Qy	258	AAGAAATCGCGAAGCCAGAGGTTACTCTTATGTGGACATTTGTGATGTGTGCCAGAG	317
Db	6041	AAGACCAAGCTCAGAGCCCACTGCTATCTCCACTATGTGTTTGAAGGAGTAAGGAACCAAGCA	6100
Qy	318	AAGCTGCCAAATTTATCAAACTCAGATCAAGAGTTCTTTTGAAGAACACCTCGCATACCGAT	377
Db	6101	GTCCCTCACAAGAAGTATGCCAGACTCAAAACAGACTTTTGAAGAGCAANTCTCTCTAAG	6160
Qy	378	GAAGAATATCGCTATTGTCTTCAAGGAGTGATACTTTGATGTGAGAGACACAAATGAT	437
Db	6161	TATGTAGGGAACAAGATCACTCAGGTGGATGAGTACATCAAGAGGCAGTGGACCATTAT	6220
Qy	438	CAGTGGATCGTATAGCACTGAAGA	462
Db	6221	GCTGGACAACCTTATGTGCTGGATA	6245

[illegible]


```

PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/550,804
FILING DATE: 09-JULY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/147,781
FILING DATE: 25-JANUARY-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/078,538
FILING DATE: 28-JULY-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 06/891,529
FILING DATE: 31-JULY-1986
ATTORNEY/AGENT INFORMATION:
NAME: Rae-Venter, Barbara
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: CGNE-037/04US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 926-6200
TELEFAX: (415) 854-3713
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4325 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FRAGMENT TYPE: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 2152..2703
US-08-453-924-2

Query Match          3.6%; Score 31.4; DB 1; Length 4325;
Best Local Similarity 64.4%; Pred. No. 8.5;
Matches 47; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 44 CTCCTGTTCCACCGCCGCATTCACATGGAGACAGCAATTCACAGATGTGAACGCCGAGGT 103
   ||| |||| | ||| |||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 4234 CCCTACTTCAATGCCCAATCTACTTTGGAAACAACAGACACAGATTTGGGAAAATTGATGA 4293
   ||| |||| | ||| |||| | ||| |||| | ||| |||| | ||| |||| | ||| ||||

QY 104 GATAGAAGCATGG 116
   ||| |||| |||| |||
DB 4294 GATCCAGCTTGG 4306

RESULT 13
US-08-471-791-29
Sequence 29, Application US/08471791
Patent No. 5723595
GENERAL INFORMATION:
APPLICANT: Thompson, Gregory A
APPLICANT: Knaut, Vic C
TITLE OF INVENTION: Plant Desaturases-Compositions
TITLE OF INVENTION: and Uses
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: California
COUNTRY: USA
ZIP: 95616
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: MicrosoftWord 5.1 (a)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,791
FILING DATE: 6-JUNE-95
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/762,762

```

```
; FILING DATE: 16-SEPT-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01746
; FILING DATE: 14-MAR-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/615,784
; FILING DATE: 14-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/567,373
; FILING DATE: 13-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/494,106
; FILING DATE: 16-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Lassen, Elizabeth
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 69-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; TELEX: 350370 CGNE
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4325 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; US-08-471-791-29

Query Match 3.6%; Score 31.4; DB 1; Length 4325;
Best Local Similarity 64.4%; Pred. No. 8.5;
Matches 47; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 44 CTCCTGCTTCCACCCGCAATCCACATGAGAACGAAATCCAGGATGCTAAGACGGAGGT 103
DB 4234 CCCTTACTTCAATGCCCACTACTTGGAAACAAGACACAGATTGGGAAAGTTGATGA 4293

QY 104 GATAGAAGCATGG 116
DB 4294 GATCCAAGCTTGG 4306

RESULT 14
PCT-US91-01746-29
; Sequence 29, Application PC/TUS9101746
; GENERAL INFORMATION:
; APPLICANT: Thompson, Gregory A
; APPLICANT: Knauf, Vic C
; TITLE OF INVENTION: Plant Desaturases-Compositions and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: California
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.7
; SOFTWARE: Microsoftword 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01746
; FILING DATE: 19910314
; CLASSIFICATION: 435
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/615,784
; FILING DATE: 14-NOV-1990
; APPLICATION NUMBER: 07/567,373
; FILING DATE: 13-AUG-1990
; APPLICATION NUMBER: 07/494,106
; FILING DATE: 16-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Lassen, Elizabeth
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; REFERENCE/DOCKET NUMBER: CGNE 69-3 WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; TELEX: 350370 CGNE
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4325 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; PCT-US91-01746-29

Query Match 3.6%; Score 31.4; DB 5; Length 4325;
Best Local Similarity 64.4%; Pred. No. 8.5;
Matches 47; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 44 CTCCTGCTTCCACCCGCAATCCACATGAGAACGAAATCCAGGATGCTAAGACGGAGGT 103
DB 4234 CCCTTACTTCAATGCCCACTACTTGGAAACAAGACACAGATTGGGAAAGTTGATGA 4293

QY 104 GATAGAAGCATGG 116
DB 4294 GATCCAAGCTTGG 4306

RESULT 15
US-09-434-408-3
; Sequence 3, Application US/09434408
; Patent No. 6440697
; GENERAL INFORMATION:
; APPLICANT: Venezia, Domenick
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: RING FINGER PROTEIN ZAPOP3
; FILE REFERENCE: 98-41
; CURRENT APPLICATION NUMBER: US/09/434,408
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: US 60/108,258
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate polynucleotide sequence of zapop3
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2169)
; OTHER INFORMATION: n = A,T,C or G
; US-09-434-408-3

Query Match 3.6%; Score 31.2; DB 4; Length 2169;
Best Local Similarity 28.4%; Pred. No. 7;
Matches 48; Conservative 36; Mismatches 85; Indels 0; Gaps 0;

QY 304 ATGTGTGCCCCAGAGAAGCTGCCAAATATGAAATCAAGATCAAGAGTTCTTTTGAAGAAC 363
DB 923 ARGGNYTNWSNGARCAYCARMGNCAYTYNGAYGNGMNCARMGNTNCARGARCAAY 997
```

364 ACCTGCATACCGATGAAGAATACGCTATTGCTTGAAGGAGTGGATFACCTTTGATGTGA 423
983 TNAARCARACNGARCARAAYATHWSNWSNMGNATHCARAARYTNTNCARGAAYAYCARM 1042
424 GAGACCAAAATGATCAGTTCGTTAGTACTGAAGAAAGGAGGCAT 472
1043 GNCARAAARWSNWSNGARATHYTHAARWSNYTNGARAAAYGARWGNAT 1091

earch completed: December 1, 2002, 21:42:13
ob time : 116 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

DM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 21:04:18 ; Search time 105 Seconds
(without alignments)

3198.292 Million cell updates/sec

Title: US-09-785-738A-1

Perfect score: 872

Sequence: 1 agacgaacacacacagaat.....acaggagtcagctgaa 872

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 341543 seqs, 192557720 residues

Total number of hits satisfying chosen parameters: 603086

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	872	100.0	872	10	US-09-785-738A-1
2	424.4	48.7	980	10	US-09-785-738A-3
3	339.2	38.9	933	10	US-09-785-738A-11
4	311.8	35.8	774	10	US-09-785-738A-5
5	307	35.2	849	10	US-09-770-445-629
6	285.6	32.8	889	10	US-09-785-738A-9
7	244	28.0	743	10	US-09-770-149-44
8	242	27.8	603	10	US-09-785-738A-7
9	213	24.4	2823	9	US-09-938-842A-1843
10	207.8	23.8	972	10	US-09-785-738A-15
11	201	23.1	1176	10	US-09-945-301-545
12	199.8	22.9	1107	10	US-09-840-787-77
13	182.8	21.0	403	10	US-09-878-574-3886
14	182.4	20.9	706	10	US-09-785-738A-17
15	181.8	20.8	920	10	US-09-785-738A-13
16	172	19.7	1093	9	US-10-191-322-1
17	154.8	17.8	456	10	US-09-880-107-1349
18	136.8	15.7	300	9	US-09-964-899-36
19	123.8	14.2	267	10	US-09-923-876-5392

ALIGNMENTS

RESULT 1

US-09-785-738A-1

; Sequence 1, Application US/09785738A

; Patent No. US2002032918A1

; GENERAL INFORMATION:

; APPLICANT: Sauter, Margaret M

; APPLICANT: Loriecke, Rene

; TITLE OF INVENTION: ALTERATION OF GROWTH AND ADAPTATION UNDER HYPOXIC

; FILE REFERENCE: 2283/201

; CURRENT APPLICATION NUMBER: US/09/785,738A

; PRIOR FILING DATE: 2001-02-15

; PRIOR APPLICATION NUMBER: 60/183,572

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 872

; TYPE: DNA

; ORGANISM: Rice

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (69)..(668)

US-09-785-738A-1

Query Match 100.0%; Score 872; DB 10; Length 872;
Best Local Similarity 100.0%; Pred. No. 3.2e-252;
Matches 872; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGACGAACAAAACAGAAATCCATCCGCAATATCGAAGGTCGCTCTTTGCTTCCACCCCG 60

Db 1 AGACGAACAAAACAGAAATCCATCCGCAATATCGAAGGTCGCTCTTTGCTTCCACCCCG 60

Qy 61 CAATCCACATGGAGACGAAATTCAGGAGGTGTAAGACGAGGTGATAGACGATGGTACA 120

Db 61 CAATCCACATGGAGACGAAATTCAGGAGGTGTAAGACGAGGTGATAGACGATGGTACA 120

Qy 121 TGGATGATAGCGAAGAGGACGAGGCTTCCTCATCACCGCGAACCACCAAGAAATTCATTC 180

Db 121 TGGATGATAGCGAAGAGGACGAGGCTTCCTCATCACCGCGAACCACCAAGAAATTCATTC 180

Qy 181 CTGTTGATAGCTTACAGAACTAGGAGTAATCAGCTGGCGCCCTAAATCTGTATACTGGG 240

Sequence 6162, Ap
Sequence 174, App
Sequence 5315, Ap
Sequence 14604, A
Sequence 81, Appl
Sequence 12507, A
Sequence 6204, Ap
Sequence 14619, A
Sequence 4750, Ap
Sequence 41, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1965, Ap
Sequence 6, Appl
Sequence 13768, A
Sequence 30307, A
Sequence 611, App
Sequence 342, App
Sequence 3897, Ap
Sequence 298, App
Sequence 9803, Ap
Sequence 898, App
Sequence 248, App
Sequence 8596, Ap

Db 181 CTTGTTGATAAGCTTACAGAACTAGGAGTAATACAGCTGCGCCTAAATCTCTGATAAAGTGGG 240
QY 241 AGAATTCGAGAACTGAGAGAAATCCGGAAGCCAGAGGTACTCTTATGTGGACATTT 300
Db 241 AGAATTCGAGAACTGAGAGAAATCCGGAAGCCAGAGGTACTCTTATGTGGACATTT 300
QY 301 GTGATGTGTGCCAGAGAGCTGCCAAATATGAACCTAAGATCAAGAGTTCTTTGAAG 360
Db 301 GTGATGTGTGCCAGAGAGCTGCCAAATATGAACCTAAGATCAAGAGTTCTTTGAAG 360
QY 361 AACCTCGATACCGATGAAGAAATACGCTATTGCTTGAAGGAGTGGATCTTTGATG 420
Db 361 AACCTCGATACCGATGAAGAAATACGCTATTGCTTGAAGGAGTGGATCTTTGATG 420
QY 421 TGAGAGCAAAATGATCAGTGGATTCGTATAGCACTGAAGAAAGGAGGCGATGATTGTTTC 480
Db 421 TGAGAGCAAAATGATCAGTGGATTCGTATAGCACTGAAGAAAGGAGGCGATGATTGTTTC 480
QY 481 TGCTTCAGGGATGTACACCGCTTTAGCTTGGACACCGACCACTATATCAAGGCAATGC 540
Db 481 TGCTTCAGGGATGTACACCGCTTTAGCTTGGACACCGACCACTATATCAAGGCAATGC 540
QY 541 GACTGTTTGTGGCGATCCTGTTTGGACACCGCTACAAACCGTCCCGCATGACCAATCTTCCCTG 600
Db 541 GACTGTTTGTGGCGATCCTGTTTGGACACCGCTACAAACCGTCCCGCATGACCAATCTTCCCTG 600
QY 601 CAAGAAAGAGTTTGTGGCTTAACTTCTCAAGTCAGAGGTGAAGTCAAGCACTTGAAG 660
Db 601 CAAGAAAGAGTTTGTGGCTTAACTTCTCAAGTCAGAGGTGAAGTCAAGCACTTGAAG 660
QY 661 GCTTCTGAGGGTTTGTGGCTTAACTTCTCAAGTCAGAGGTGAAGTCAAGCACTTGAAG 720
Db 661 GCTTCTGAGGGTTTGTGGCTTAACTTCTCAAGTCAGAGGTGAAGTCAAGCACTTGAAG 720
QY 721 TGCATAGCAATGTAATTTAGCACAGTGGCTATGCTGGCCACTCACCACCTTGAAGTGA 780
Db 721 TGCATAGCAATGTAATTTAGCACAGTGGCTATGCTGGCCACTCACCACCTTGAAGTGA 780
QY 781 AAGATTTAATGATTTTGTAACTTATGATCAATCGGCATATAGCAATTCGGAATG 840
Db 781 AAGATTTAATGATTTTGTAACTTATGATCAATCGGCATATAGCAATTCGGAATG 840
QY 841 TGTTCATTAACAGGAGTCAAGCACTGAA 872
Db 841 TGTTCATTAACAGGAGTCAAGCACTGAA 872

RESULT 2
US-09-785-738A-3
; Sequence 3, Application US/09785738A
; Patent No. US20020032918A1
; GENERAL INFORMATION:
; APPLICANT: Sauter, Margret M
; APPLICANT: Lorbiecke, Rene
; TITLE OF INVENTION: ALTERATION OF GROWTH AND ADAPTATION UNDER HYPOXIC
; FILE REFERENCE: 2283/201
; CURRENT APPLICATION NUMBER: US/09/785,738A
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,572
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 980
; TYPE: DNA
; ORGANISM: Rice
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (139)..(735)
US-09-785-738A-3

Query Match 48.7%; Score 424.4; DB 10; Length 980;
Best Local Similarity 83.3%; Pred. No. 1.1e-117;
Matches 495; Conservative 0; Mismatches 96; Indels 3; Gaps 1;
QY 67 ACATGAGAACGAAATCCAGGATGGTAAGACCGAGGTGATAGAACGATGGTACATGATG 126
Db 137 AGATGGAGAACCAATTCAGGATGGCAAGGAGGAGGTGATCGAAGCTTGGTACATGATG 196
QY 127 ATAGCGAAGAGGAGGAGGCTTCCTCATCACCGGCAACCAAGAAATTCATTCCTGTG 186
Db 197 ACATGAGAGGAGGAGGCTTCCTCATCATCGTGAGCCCAAGAAATTCATTCCTT 256
QY 187 ATAGCTTACAGAACTAGGAGTAATCAGCTGGCGCTTAATCTCTGATTAACCTGGAGAA 246
Db 257 GCAAACTTTTCAGAGTTAGGAATATTAAAGCTGGCGCTGAATGCTGATGACTGGGAGATG 316
QY 247 GCAGAACCTTGAAGAGAAATCCGGAAGCCAGAGGTTACTCTTATGTGGACATTTGTGATG 306
Db 317 ATGAGAACCTCAAGAAATCCGCTGAGCCGAGGAGTACTCTTACATGATATTTGTGATG 376
QY 307 TGTGCCAGAGAGCTGCCAAATATGAACTAAGATCAAGAGTTTCTTTGAAGAACACC 366
Db 377 TGTGTCCAGAAAGCTGCCAAACTATGAGGCTTAAGCTGAAATTTCTTTGAAGAACACT 436
QY 367 TGCATACCGATGAAGAAATACGCTATTGCTTGAAGGAGTGGATACTTTGATGTGAGAG 426
Db 437 TGCATACTGATGAAGAGATACGCTATTGCTTGAAGGAGTGGATACTTCGATGTGAGG 496
QY 427 ACCAAATGATCAGTGGATTCGTATAGCACTGAAGAAAGGAGGAGTATTTCTTGCCTG 486
Db 497 ACCAAATGATCAGTGGATTCGTATAGCACTGAAGAAAGGAGGAGTATTTCTTTGCTG 556
QY 487 CAGGATGTACACCGCTTTAGCTTGGACACCGCAACTATATCAAGCAATGCGACTGT 546
Db 557 CGGATGTATACACCGCTTCACATTTGGACAGTGAACAATCAAGCAATGCGGCTCT 616
QY 547 TTGTGGCGATCCTGTTTGGACACCGTACACCGTCCCATGACCATCTTCTCTCAAGAA 606
Db 617 TTGTGGGAGAGCTGTCTGGAGCGCTACACCGTCCCATGACCATCTGCGAGTAGAA 676
QY 607 AGAGTTTGTGGCTAACTTCTCAAGTCAGAGGTGAAGTGAAGTGAAGTGAAG 660
Db 677 AGAGTATGTCGAAAAAATATCAA---CAGGGGTGGAACTCAAGCTGTCGAAG 727

RESULT 3
US-09-785-738A-11
; Sequence 11, Application US/09785738A
; Patent No. US20020032918A1
; GENERAL INFORMATION:
; APPLICANT: Sauter, Margret M
; APPLICANT: Lorbiecke, Rene
; TITLE OF INVENTION: ALTERATION OF GROWTH AND ADAPTATION UNDER HYPOXIC
; FILE REFERENCE: 2283/201
; CURRENT APPLICATION NUMBER: US/09/785,738A
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,572
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Cotton
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (33)..(635)
US-09-785-738A-11
Query Match 38.9%; Score 339.2; DB 10; Length 933;
Best Local Similarity 74.8%; Pred. No. 4.2e-92;
Matches 425; Conservative 0; Mismatches 143; Indels 0; Gaps 0

[illegible][illegible]

Db	18	GGATGTCATACAGGCATGGTACATGATGACAAACGATGAGGACCAGAGCGCTTCCTCATCA	77
Qy	158	CCGCGAACCCAAAGAATTCACTCTCTTGAAGCTTACAGAACTAGGAGTAATCAGCTG	217
Db	78	CCGTGAGCCAAAGAAATTTGCTGCTCTTGCACAGCTGGCTGAACTTGGAGTGCTCAGCTG	137
Qy	218	GCGGCTAAATCCTGATAACTGGGAGAAATTCGCGAARACCTTGAGAGAAATCCGCGGAAGCCAG	277
Db	138	GAGACTTGATGCTGACAATTTATGAGACTGATGAGGAGTTGCAAGAAATTCGGGAAGATCG	197
Qy	278	AGGTTACTCTTATGTGGACATTTGTGATGTGTGCCAGAGAAAGCTGCCAAATATATGAAC	337
Db	198	TGATATTCAATCATGATTGTTCTGTGAGGTTGCCCTGAGAACTACCGAATACGAGA	257
Qy	338	TAAGATCAAGAGTTCTTTTGAAGAACACCTGCATACCGATGAGAGAAATACGCTATTGCT	397
Db	258	GAATAACAAGAACTTTTGAAGAACACCTGCACACCGAGGAGATCCGTTACGCTGT	317
Qy	398	TGANGGAGTGGATACTTTGATGTGAGAGACCAAAATGATCAGTGGATTCGTTATAGCACT	457
Db	318	TGCAGGAAGTGGTTACTTTTGTATGTCCGCGATGTGAATGAGAGCTGGATTCGGGCTGGGT	377
Qy	458	GAAGAAAGGAGGCATGATTTCTTGCTCGCAGGAGTGTACACCGCTTTACGTTTGSACAC	517
Db	378	AAAGAAAGTGGATGATTGTTCTTCTGCTGGAATCTATCACCGCTTCACGCTTGATTC	437
Qy	518	CGACAACATATACAAGCAATGCGACTGTFTTGTGGCGATCTCTGTTTGGACACCCCTACAA	577
Db	438	AAGCAACTACATTAAGGCAATGCGTCTCTTTGTTGGTGACCCCAATTTGGACTCCATACAA	497
Qy	578	CCGTGCCCATGACCATCTTCTCGCAAGAGAGTTTGTGGCTAAACTTCTCAAGTCAGA	637
Db	498	TCGTCCACATGATCATCTCCCGCAGGCAAGAAATATGTTGAGACGCTTGTCAACGCAGA	557
Qy	638	AGGTGAAAAATCAAGCACTTGAAG	660
Db	558	TGGCGCTGGTCGTCTGTAAATG	580

```

RESULT 5
US-09-770-445-629/c
; Sequence 629, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; TITLE OF INVENTION: thaliana
; FILE REFERENCE: 20230US (PARA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 629

```

;> LENGTH: 849
;> TYPE: DNA
;> ORGANISM: Arabidopsis thaliana
;> FEATURE:
;> NAME/KEY: misc_feature
;> LOCATION: (1)...(849)
;> OTHER INFORMATION: n = A,T,C or G
US-09-770-445-629

Query Match 35.2%; Score 307; DB 10; Length 849;
Best Local Similarity 72.4%; Pred. No. 1.9e-82;
Matches 391; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
QY 83 CCAGATGGTAAGCGGAGGTGATAGAGCATGGTACATGGATGATAGCGAGAGACCA 142
Db 849 CAAGGATGGGAGAGAGAGAGTTCAGCTTGGTATATGATGATAGTCAAGAGATCA 790
QY 143 GAGGCTTCCTCATCACCAGCGAAGCAAGCAATTCCTCTGTTGATAAGCTTACAGAACT 202
Db 789 GAGACTTCTCACCACAGGATCCTAAGAGCTTGTATCGTTGGGCAAACTTGCAGAGCT 730
QY 203 AGAGTAATACAGTGGCGCCCTAAATCCTGTATACCTGGGAGAAATTCGAGAACTGAAGAG 262
Db 729 GGGAGTACTTACGTGGAGCTTGATGCTGATACCTAGCAAAACCATGAGGATTTGAAAA 670
QY 263 ANTCCGGAAGCCAGAGCTTACTCTTATGTGGACATTTGTGATGTGTCGCCAGAGAGCT 322
Db 669 GATCCGTAATCTCGTGTACTCTTACATGGAGCTTTGTGAGGTATGCCCGGAAAGCT 610
QY 323 GCCAATATGAACACTAAGATCAAGAGTCTTTTGAAGAACACCTCATACCGATGAAGA 382
Db 609 TCCAAACTATGAAGTGAAGTAAAGAGCTTTTCGAAGAACATTTGCATACATGAAGA 550
QY 383 AATACGCTATTGCTTGAAGGAGTGGATCTTATGTGGACATTTGTGATGTGTCGCCAGAGCT 442
Db 549 GATCCGTTACTGCTTGCAGAGAGTGGTACTTTGATGTGAGAGATCGCAATGAAGCTTG 490
QY 443 GATTCGTATAGCACTGAAGAAGGAGCATGATGTTCTGCTCGAGGAGTGTACCAACG 502
Db 489 GATTAGGCTATGGGTAAANNNNNAGCTATGATANNNTTACCTGCTGGGATCTATCATCG 430
QY 503 CTTTACGTTGACACCGACACTATATCAAGGCAATCGGACTGTTTCTTGGCGATCCTGT 562
Db 429 CTTACCGTGGACTCTGACAACTATATCAAGGCAATCGGCTTTTCTGGGTGAACCGGT 370
QY 563 TTGGACCCCTACACCGTCCCATGACCATCTTCTCTGCAAGAAAGAGTTTTTGGCTAA 622
Db 369 ATGGACACCATACAATCGCCCAACAGACCATCTTCTCTGCAAGAAAGATATGTCGATA 310

RESULT 6
US-09-785-738a-9
; Sequence 9, Application US/09785738a
; Patent No. US2002032918A1
; GENERAL INFORMATION:
; APPLICANT: Sauter, Margaret M
; APPLICANT: Lorbiecke, Rene
; TITLE OF INVENTION: ALTERATION OF GROWTH AND ADAPTATION UNDER HYPOXIC
; FILE REFERENCE: 2283/201
; CURRENT APPLICATION NUMBER: US/09/785,738A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,572
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 9
; LENGTH: 889
; TYPE: DNA
; ORGANISM: Soybean
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(634)

US-09-785-738a-9

Query Match 32.8%; Score 285.6; DB 10; Length 889;
Best Local Similarity 68.3%; Pred. No. 5.2e-76;
Matches 396; Conservative 0; Mismatches 184; Indels 0; Gaps 0;
QY 83 CCAGATGGTAAGCGGAGGTGATAGAGCATGGTACATGGATGATAGCGAGAGACCA 142
Db 46 CAAGGATCCACGAGAGAGTCTCTTCAAGCCTGGTACATGATGATAGTGAAGATCA 105
QY 143 GAGGCTTCCTCATCACCAGCGAAGCAAGCAATTCCTCTGTTGATAAGCTTACAGAACT 202
Db 106 AGACTCCCCACCAAGAGAGGAGTTGTCTGTTGGACCACTTGTCTGAAGT 165
QY 203 AGAGTAATCAGCTGGCGCTAAATCCTGTATACCTGGGAGAAATTCGAGAACTGAAGAG 262
Db 166 TGGAGTCTTAGCTGGAACCTAGATCTGATAACCATGAANAATGATCCAGAGCTGAAGAA 225
QY 263 ANTCCGGAAGCCAGAGGTACTCTTATGTGGACATTTGTGATGTGTCGCCAGAGAGCT 322
Db 226 GATTCGTTGAAGAGCGGTTCACACTTACATGGATGTTGTGAGGTCTGCCAGAAAAGTT 285
QY 323 GCCAATATGAACACTAAGATCAAGAGTCTTTTGAAGAACACCTGTCATACCGATGAAGA 382
Db 286 GCCAATATGAACAGAAAATCAAAAGCTTCTTTGAAGAGCATCTTCACACTGATGAGGA 345
QY 383 AATACGCTATTGCTTGAAGGAGTGGATCTTGTGATGTGAGAGACCAAAATGATCAGTG 442
Db 346 GATCCGCTTTTGTGCTGCTGGAAGTGGCTATTTGTGATGTAGGATCGCAATGAAGCTTG 405
QY 443 GATTCGTATAGCACTGAAGAAGGAGCATGATGTTCTGCTCGAGGAGTGTACCAACG 502
Db 406 GATTCGTGTGCTGCTCAAGAAAGGAGGATGATCATCTTACCTGCCGAATTTATCATCG 465
QY 503 CTTTACGTTGGACACCGACAACTATATCAAGGCAATGGAGCTGTTTGTGGCGATCCTGT 562
Db 466 CTTTACGCTAGTATGAGAGCAACTACATTAAGGCTTTGAGGCTTTTGTGTGAGGCCAGT 525
QY 563 TTGGACACCCCTACAAACGCTCCCATGACCATCTTCTCTGCAAGAAAGAGTTTTTGGCTAA 622
Db 526 TTGAGCTCCATACATAATCGTCCAAATGACCATCTCCTCGAAGACAAATATGTCAGGA 585
QY 623 ACTTCTCAAGTCAGAGGTGAAATAACACGACTTTGAAGGC 662
Db 586 TTTTGTGGAAGAGATGTTAGCAGCCATCTGTTGATGCC 625

RESULT 7
US-09-770-149-44
; Sequence 44, Application US/09770149
; Patent No. US2002005963A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; thaliana

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```
; LENGTH: 2823
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1843

Query Match      24.4%; Score 213; DB 9; Length 2823;
Best Local Similarity 62.8%; Pred. No. 6.1e-54;
Matches 400; Conservative 0; Mismatches 150; Indels 87; Gaps 1;

QY 69 ATGGAGACGAATTCACAGATGGTAAGACGGAGGTGATAGAACATGGTACATGGATGAT 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 ATGGGTGAAGCGGTCAAGGATGGGAGAGAGAGTATTCAGCTGGTATATGGATGAT 60

QY 129 ACCGAAGAGGACCAAGGCTTCCTCATCACCGGAGACCCAAAGAAATTCCTCGTTGAT 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AGTGAAGAGGATCAGAGACTTCCACCAACAAGATCCTAAAGAGTTGTATCGTTGGAC 120

QY 189 AAGCTTACAGAACTAGGAGTAAATCAGCTGGCGCTAAATCCTGATACTGGGAGAAATTC 248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 AAACCTGCAGAGCTGGGAGTACTTAGCTGGAGACTTGATCGCTAATACTATGAACCGAT 180

QY 249 GAGAACCTGAAGAGAAATCCGCGAAGCCAGAGGTTACTCTTATCT 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 GAGGATTTGAAAGATCCGTTGAATCTCGTGGTTACTCTTCATCGTTGTGGTCTTTAT 240

QY 293 ----- 292

Db 241 TTATTACTATCCCACTAGACTTGAATGGTAGATGTTTATTTTGTCCAAGTCTTATTT 300

QY 293 -----GGACATTTGATGTGTGCCAGAGAGCTGCCAAATTAAGAACTAAG 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 AACTCTTAGGGACTTTGTGAGGTATGCCGGAAGATTCACAACTATGAAGTGAA 360

QY 342 ATCAAGATTTCTTTGAAGACACCTGCATACCGATGACGAAGAAATACGCTATTGCTTGA 401
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 GTGAAGAGCTTTTCGAAGACATTTACACACTGATGAGGAGATCGGTTACTGCGTTGCA 420

QY 402 GGGAGTGATCTTTGATGTGAGAGACCAAAATGATCAGTGGATTCGTATACACTGAAG 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 GGAATGTTACTTTGATGTGAGAGATCGTAATGAAGCTTGGATAGGGTATGGTAAAG 480

QY 462 AAAGGAGCATGATTTCTCCCTGCAGGATGTACACCGCTTTACGTTGGACACCGAC 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 AAGGGAGGTATGATAGTCTTACCTGCTGGGATCTATCATCGCTTCACTGCGACTCTGAC 540

QY 522 AACTATATCAAGCAATCGACTCTTTCTTGGCGATCCTGTTTGGACACCCCTACACCGT 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 AACTATATCAAGCAATCGGCTATTCTGGGTGAACCGGTATGGACACCATACAATCCG 600

QY 582 CCCATGACCATCTTCTTCGCAAGAAAGAGTTTGTG 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 CCACAGCACCATCTTCTCGCAAGATGCTTGTGTTGG 637
```

```
RESULT 10
US-09-785-738A-15
; Sequence 15, Application US/09785738A
; Patent No. US20020032918A1
; GENERAL INFORMATION:
; APPLICANT: Sauter, Margret M
; APPLICANT: Lorbiecke, Rene
; TITLE OF INVENTION: ALTERATION OF GROWTH AND ADAPTATION UNDER HYPOXIC
; FILE REFERENCE: 2283/201
; CURRENT APPLICATION NUMBER: US/09/785,738A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,572
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 972
; TYPE: DNA
```

```
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17)..(556)
US-09-785-738A-15

Query Match      23.8%; Score 207.8; DB 10; Length 972;
Best Local Similarity 63.5%; Pred. No. 1.3e-52;
Matches 317; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 103 TGATAGAAGCATGTGTACATGTGATAGCAGAGAGGACGAGAGGCTTCTCATCACCGCG 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18 TGGTCAGCGCTGGTATATGGAGAGTCCACCGCGGCCCGCGGAAGCCACCGCGCAC 77

QY 163 AACCCAAAGAAATTCATCTCTGTGTAGACCTTACAGAACTAGGAGTAACTACAGTGGCGCC 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 AGCCCGACCGCCCGGTGAGCCTGGAGCAGCTGCGCACGCTCGGAGTGCCTTATTGGAGC 137

QY 223 TAAATCCTGATAACTGGGAGAAATTCGAGAACTTGAAGAGAAATCCGGAAGCCAGAGGTT 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 TAGATGCTGACAAGTATGAGAACGATCCAGAACTAGAAAAGATCCGGAAGAAATGAGAACT 197

QY 283 ACTCTTATGTGACATTTGTGATGTGTGCCCGAGAACTGCCAAATTAAGAACTAAGA 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 198 ACTCTGGGATGGACATCATCACCATATGCAAGATACACTTCCCAATTTACGAGGAGAA 257

QY 343 TCAAGAGTTTCTTTGAAGAACACCTGCATACCGATGAAGAAATACGCTATTGTCTTGAAG 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 TCAAGATGTTCTTTGAGGAACATCTGCATCTCGATGAGGAGATCCGCTTACATCTCGAGG 317

QY 403 GGAGTGGATCTTTGATGTGACAGACCAAAATGATCAGTGGATTCGTATAGCACTGAAGA 462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 GTAGTGGGTACTTCGATGTCCAGGACAAGGAGCAAGTGGATTCGGATTTCCATGGAGA 377

QY 463 AAGGAGGCATGATTTGTTCTGCTGCGAGGATGTACCCGCTTTACGTTGGACACCGACA 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 AGGGGACATGATTACTCTCTCCGCGCATCTATACCGCTTCACACTGAGCAGAGAAGA 437

QY 523 ACTATATCAAGCAATGCGACTGTTTGTGGCGATCCTGTTTGGACACCCCTACACCGCTC 582
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 ATTACGTGAAGGCCATGCGGCTGTTTGTGGAGAACCTGTGTGGACACCATACAACCGGC 497

QY 583 CCCATGACCATCTTCTCTGC 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 CAGTGACCATTTTGATGC 516
```

```
RESULT 11
US-09-925-301-545
; Sequence 545, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 545
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-545

Query Match      23.1%; Score 201; DB 10; Length 1176;
Best Local Similarity 61.7%; Pred. No. 1.6e-50;
Matches 318; Conservative 1; Mismatches 196; Indels 0; Gaps 0;
```


[illegible]

```

381 GAAATACGCTATTGTGTAAGGAGTGGATCTTTGATGTGAGACCAAAATGATCAG 440
  || || |||||  || || || ||||| ||||| ||||| ||| | |||||
304 GAGATCCGCTACATCTGGATGGCAGTGGTACTTCGATGTGAGGACAAGGAGGACCAG 363
  ||||| || || ||||| ||||| ||||| ||||| ||||| |||||
441 TGGATTCGTATAGCACTGAAGAAAGGAGGCATGATGTTCTGCTCGAGGGATGTACCAC 500
  ||||| || || ||||| ||||| ||||| ||||| ||||| |||||
364 TGGATCCGGATCTTCATGGAGAGGGAGACATGTTGACGCTCCCGGGGATCTATCAC 423
  ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
501 CGCTTTACGTTGGACACCGACAATATCAAGGCAATGCGACTGTTTGTGGCGATCCT 560
  ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
424 CGCTTCACCGTGGACGAGGAAGAACTACACGAAGGCCATGCGGCTGTTGTGGGAGAACCG 483
  ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
561 GTTGGACACCCCTACAACCGTCCCATGACCATCTTCTGCAAGAAAGGAGTTTTGG 617
  ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
  484 GTGTGGACAGGTACAACCGCGCGCTGACCATTTTGAAGCCCGCGGGCAGTACGTG 540

```

Search completed: December 1, 2002, 22:49:06
 Job time : 112 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 20:00:28 ; Search time 2187 Seconds

(without alignments)
6457,460 Million cell updates/sec

Title: US-09-785-738a-1

Perfect score: 872
Sequence: 1 agacgaacacacacagat.....acagagatcatgaagctgaa 872

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_juv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	533.2	61.1	608	9	AU100980 AU100980
2	456.4	52.3	618	13	BM324741 PIC1_35_A
3	439.4	50.4	591	13	BM324752 PIC1_35_B
4	438	50.2	546	10	AM284744 LG1_214_H
5	421.6	48.3	700	9	AL504809
6	419.8	48.1	1097	11	AY103746

7	419.6	48.1	467	14	C28663	C28663	C28663 Rice
8	419.4	48.1	729	14	B0806508	B0806508	B0806508 WHE3579_H
9	418	47.9	574	10	AM678360	WS1_15_C1	AM678360 WS1_15_C1
10	417.8	47.9	600	13	B1542476		B1542476 949020B04
11	417.8	47.9	978	12	BG343468	HVSM8000	BG343468 HVSM8000
12	417.2	47.8	642	10	AV932455		AV932455 AV932455
13	417.2	47.8	642	13	BJ471108		BJ471108 BJ471108
14	417.2	47.8	644	14	BQ464906	HU02B15t	BQ464906 HU02B15t
15	417.2	47.8	709	13	BM370338	EBRC08_SQ	BM370338 EBRC08_SQ
16	417.2	47.8	840	12	BF6292917	HVSM8001	BF6292917 HVSM8001
17	417.2	47.8	890	14	BM816671	HC114G05_	BM816671 HC114G05_
18	416.6	47.8	605	14	BQ244802	TAE15035C	BQ244802 TAE15035C
19	416	47.7	715	14	B0805282	WH3565_A	B0805282 WH3565_A
20	415.6	47.7	697	14	BQ755512	EBed07_SO	BQ755512 EBEd07_SO
21	415.6	47.7	794	12	BG299990	HVSM8002	BG299990 HVSM8002
22	414	47.5	563	10	AM126415	614071A04	AM126415 614071A04
23	414	47.5	693	10	BE215135	HV_CED000	BE215135 HV_CED000
24	412.4	47.3	793	12	BG367078	HVSM1001	BG367078 HVSM1001
25	411.2	47.2	611	14	BQ466283	HT02B17r	BQ466283 HT02B17r
26	410.8	47.1	802	10	BE195423	HVSM8008	BE195423 HVSM8008
27	402.8	46.2	662	9	AL508895		AL508895 AL508895
28	401.4	46.0	560	9	AI947785	603028D09	AI947785 603028D09
29	399.6	45.8	583	14	BQ758819	EBna07_SO	BQ758819 EBna07_SO
30	397.8	45.6	566	10	AV915690	AV915690	AV915690 AV915690
31	396.8	45.5	589	10	AV941388	AV941388	AV941388 AV941388
32	396.8	45.5	626	10	AV922964	AV922964	AV922964 AV922964
33	396.6	45.5	549	14	BQ249643	TAE25015G	BQ249643 TAE25015G
34	393.4	45.1	548	14	BQ238245	TAE05006E	BQ238245 TAE05006E
35	393.2	45.1	713	12	BG463642	EM1_50_B1	BG463642 EM1_50_B1
36	390.8	44.8	584	9	AJ432173	AJ432173	AJ432173 AJ432173
37	390.2	44.7	613	14	BQ238274	TAE05006B	BQ238274 TAE05006B
38	389.8	44.7	585	14	BQ458741	HA04F17r	BQ458741 HA04F17r
39	389.4	44.7	871	13	B1951424	HVSM1002	B1951424 HVSM1002
40	389.2	44.6	608	14	BQ238805	TAE05040H	BQ238805 TAE05040H
41	387.6	44.4	464	14	C27474	Rice	C27474 Rice
42	387	44.4	524	10	BE403815		BE403815
43	384.4	44.1	493	14	BQ282977	WHE3084_C	BQ282977 WHE3084_C
44	381.8	43.8	562	14	BQ758441	EBna05_SO	BQ758441 EBna05_SO
45	381	43.7	575	10	AV939451		AV939451

ALIGNMENTS

RESULT 1
AU100980
LOCUS
DEFINITION
AU100980 Rice callus Oryza sativa (japonica cultivar-group) cDNA
clone C61979, mRNA sequence.
ACCESSION
AU100980
VERSION
AU100980.1 GI:9865230
KEYWORDS
EST.
ORGANISM
Oryza sativa (japonica cultivar-group).
Oryza sativa (japonica cultivar-group).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 608)
REFERENCE
Sasaki,T. and Yamamoto,K.
Rice cDNA from callus (2000)
Unpublished (2000)
JOURNAL
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@agr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT = "RGP".
FEATURES
source
Location/Qualifiers
1..608
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"

```

/db_xref="taxon:39947"
/clone_lib="Rice callus"
/mote="Vector: pBluescript II SK+; Site.1: SalI; Site.2:
NotI; cDNA prepared from rice callus mRNAs by using
oligo(dT) as a primer and ligating to the SalI-NotI site
of pBluescript II SK+ phagemid."
BASE COUNT      198 a      111 c      131 g      166 t      2 others
ORIGIN
Query Match      61.1%; Score 533.2; DB 9; Length 608;
Best Local Similarity 97.3%; Pred. No. 8,8e-155;
Matches 541; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 310 GCCCAGAGAGCTGCCAAATTATGAACATAAGATCTTCTTGAGACACCTGC 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 GCCCANAGAGCTGCCAAATTATGAACATAAGATCTTCTTGAGACACCTGC 65
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 370 ATACCGATGAAGAAATACGCTATGTCTTGAGGGGAGGATCTTATGATGAGAGCC 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66 ATACCGATGAAGAAATACGCTATGTCTTGAGGGGAGGATCTTATGATGAGAGCC 125
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 430 AAAATGATCAGTGATTCGTATAGACATGAGAAAGAGGAGCATGTCTGCTGCAG 489
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 AAAATGATCAGTGATTCGTATAGACATGAGAAAGAGGAGCATGTCTGCTGCAG 185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 490 GGATGTACCAACCGCTTACGTTGACACGACACATATTCAGGCATGCGACTGTTG 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186 GGATGTACCAACCGCTTACGTTGACACGACACATATTCAGGCATGCGACTGTTG 245
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 550 TTGGCATCTGTTTGGACACCCCTACACCCGTCGCCATGACCATCTTCTGCAAGAAAG 609
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 246 TTGGCATCTGTTTGGACACCCCTACACCCGTCGCCATGACCATCTTCTGCAAGAAAG 305
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 610 AGTTTGGCTAAACTTCACACTCAGAGGAGGAAATTCAGCATTTAGAGCTTTGAG 669
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 306 AGTTTGGCTAAACTTCACACTCAGAGGAGGAAATTCAGCATTTAGAGCTTTGAG 365
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 670 GGTTTGTGGCTCCTGACACTGCGGTCTATATTCACCTGAATAGATGCTATAGC 729
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 366 GGTTTGTGGCTCCTGACACTGCGGTCTATATTCACCTGAATAGATGCTATAGC 425
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 730 AATGTAATTTACACAGAGGCTATGCTGCCACCTCACCACCTTAAAGTGAAGATTAA 789
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 426 AATGTAATTTACACAGAGGCTATGCTGCCACCTCACCACCTTAAAGTGAAGATTAA 485
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 790 TGATTTTGTAAATCTTATGATCAATCGCATATAGCATTTCCGAATGTGTTTCAA 849
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 486 TGATTTTGTAAATCTTATGATCAATCGCATATAGCATTTCCGAATGTGTTTCAA 545
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 850 TAAACAGAGTCAATGA 865
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 546 TAAACAGAGTCAATGA 561
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
BM324741
LOCUS      618 bp      mRNA      linear      EST 04-JAN-2002
DEFINITION bicolour cDNA, mRNA sequence.
ACCESSION  BM324741
VERSION     BM324741.1 GI:18063646
KEYWORDS   EST.
SOURCE     sorghum.
ORGANISM   Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE  1 (bases 1 to 618)
            Cordonnier-Pratt,M.-M., Gingle,A., Fang,G.C., Dean,R., Wing,R.,
            Sudman,M. and Pratt,L.H.
            An EST database from Sorghum: plants infected with a compatible
            pathogen
TITLE

```

```

JOURNAL      Unpublished (2002)
COMMENT      Contact: Cordonnier-Pratt MM
              Laboratory for Genomics and Bioinformatics
              The University of Georgia, Department of Plant Biology
              Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
              Tel: 706 542 1860
              Fax: 706 583 0210
              Email: mmp@pratt.uga.edu
              Sequences have been trimmed to exclude polyA, vector, and regions
              below Phred quality 16. The threshold for highest quality sequence
              is 20. Three-prime sequences, which are obtained with PolyTmix or
              T7 sequencing primer, are presented as the reverse complement.
              Seq primer: JEN REV
              High quality sequence stop: 604
              POLYA-No.

FEATURES
    source          location/qualifiers
    1..618
    /organism="Sorghum bicolor"
    /cultivar="Brx623"
    /db_xref="taxon:4558"
    /clone_lib="Pathogen-infected compatible 1 (PIC1)"
    /tissue="Leaves"
    /dev_stage="4-week-old seedlings infected with
    Colletotrichum graminicola"
    /note="Vector: pBluescript II SK(-) from Lambda zap II;
    Site.1: XhoI; Site.2: EcoRI; Four-week-old sorghum
    seedlings were sprayed with spore suspension prepared from
    3-week-old FRM421, a sorghum isolate of the anthracnose
    pathogen Colletotrichum graminicola. Inoculated plants
    were kept in a 25 C dark growth chamber with 100% relative
    humidity for 24 hr, followed by 12/12 hr of light/dark
    cycle at 25 C with 90% relative humidity for another 24
    hr. All leaves were harvested and quick frozen with liquid
    nitrogen and stored in a -80 C freezer. The library was
    made from poly-A RNA in the cloning vector lambda Zap II.
    Clones to be sequenced were prepared by mass excision.
    WARNING: While most or all ESTs are expected to derive
    from the host plant, no effort was made to eliminate ESTs
    deriving from the pathogen."

BASE COUNT      187 a      124 c      162 g      145 t
ORIGIN
Query Match      52.3%; Score 456.4; DB 13; Length 618;
Best Local Similarity 89.0%; Pred. No. 7.9e-131;
Matches 493; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 65 CCACATGGAGAGCAATTCAGATGTTAGACGAGGATGATGAGCATGGTACATGGA 124
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 63 CACATGGAAGAACCGAGTTCCAGATGTTAGACGAGGATGATGAGCATGGTACATGGA 122
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 125 TGATAGCGAAGAGACAGAGGCTTCTGATCAACGCCGACCAAGCAATTCATCTGCT 184
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 123 CGATAGTGAAAGAGACAGAGGCTTCTGATCAACGCCGACCAAGCAATTCATCTGCT 182
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 185 TGATAGGCTTACAGACAGATGATCACTGCGGCTTAATCCGTGATTAATCTGGAGAA 244
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 183 TGATAGGCTTACAGACAGATGATCACTGCGGCTTAATCCGTGATTAATCTGGAGAA 242
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 245 TTGCGAGACCTGGAAGAAATCCGCGAAGAGGTTACTTATGTCGACATTTGGA 304
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 243 TGATGAGAACTCAAGAAATCCGCGAAGAGGTTACTTATGTCGACATTTGGA 302
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 305 TGTGTGCCAGAGAGAGCTGCCAAATTAATGAATGATCAATGATTTCTTGAAGACA 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 303 TGTATGCCAGAGAGAGCTGCCAAATTAATGATTAATGAATGATTTCTTGAAGACA 362
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 365 CCGTCAATACCGATGAAGAAATAGCGTATGCTTGAAGAGAGTGAATCTTGAATGATGAG 424
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 363 CCGTCAATACCGATGAAGAAATAGCGTATGCTTGAAGAGAGTGAATCTTGAATGATGAG 422
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 425 AGACCAAAATGATGAGTGATGCTATAGACATGAAGAAAGACATGATGTTCTGCC 484
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 423 GACGAAATATGACGATGATCGGTGACGATGAAGAAAGACATGATGTTCTTCC 482
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

FEATURES	SOURCE
RESULT 3	
LOCUS BM324752	
DEFINITION	BM324752 591 bp mRNA linear EST 04-JAN-2002
ACCESSION	PICT_33.B12.b1_A002 Pathogen-Infected compatible 1 (PICT1) Sorghum
VERSION	BM324752
KEYWORDS	BM324752.1 GI:18063663
SOURCE	EST.
ORGANISM	sorghum.
REFERENCE	Sorghum bicolor
AUTHORS	Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta; Spermatophyta: Magnoliophyta: Liliopsida: Poales; Poaceae; PAC
TITLE	Clade: Panicoideae; Andropogoneae; Sorghum.
JOURNAL	1 (bases 1 to 591)
COMMENT	Cordonnier-Pratt,M.-M., Gingle,A., Fang,G.C., Dean,R., Wing,R., Sudman,M. and Pratt,L.H.
	An EST database from Sorghum: plants infected with a compatible pathogen
	Unpublished (2002)
	Contact: Cordonnier-Pratt MM
	Laboratory for Genomics and Bioinformatics
	The University of Georgia, Department of Plant Biology
	Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
	Tel: 706 542 1860
	Fax: 706 563 0210
	Email: mmpatt@uga.edu
	Sequences have been trimmed to exclude polyA, vector, and regions below Phred quality 16. The threshold for highest quality sequence is 20. Three-prime sequences, which are obtained with polyTmix or T7 sequencing primer, are presented as the reverse complement.
	Seq primer: JEN REV
	High quality sequence stop: 442
	POLYA=NO.
	Location/Qualifiers
	1..591
	/organism="Sorghum bicolor"
	/cultivar="BR623"
	/db_xref="taxon:4558"
	/clone_lib="Pathogen-infected compatible 1 (PICT1)"
	/tissue_type="Leaves"
	/dev_stage="4-week-old seedlings infected with
	Colletotrichum graminicola"
	/note="Vector: pBluescript II SK(-) from Lambda Zap II;
	Site_1: XhoI; Site_2: EcoRI; Four-week-old sorghum
	seedlings were sprayed with spore suspension prepared from
	3-week-old FRM421, a sorghum isolate of the anthracnose
	pathogen Colletotrichum graminicola. Inoculated plants
	were kept in a 25 C dark growth chamber with 100% relative
	humidity for 24 hr, followed by 12/12 hr of light/dark
	cycle at 25 C with 90% relative humidity for another 24
	hr. All leaves were harvested and quick frozen with liquid
	nitrogen and stored in a -80 C freezer. The library was
	made from poly-A RNA in the cloning vector lambda Zap II.
	Clones to be sequenced were prepared by mass excision.
	WARNING: While most or all ESTs are expected to derive
	from the host plant, no effort was made to eliminate ESTs
	deriving from the pathogen."
BASE COUNT	
	179 a 118 c 153 g 141 t

Query Match	50.4%	Score 439.4	DB 13	Length 591
Best Local Similarity	89.4%	Pred. No.1,5e-125		
Matches 473	Conservative 0	Mismatches 56	Indels 0	Gaps 0
QY 65	CCACATGAGAAACGAATTCACAGATGGTATAGACGAGAGTATATGAAAGCATGTGACATGGA	124		
Db 63	CAACATGGAACCGAGTTCACAGATGGTATAGACGAGAGTATATGAAAGCATGTGACATGGA	122		
QY 125	TGATATCGGAAGGAGCCAGAGAGCTTCCATCACACCGCGGAACCCAAAGAAATTCATCTCT	184		
Db 123	CGATATCTGGAAGAAAGACAGAGAGCTTCCATCACACCGCGATCCAAAGAAATTCATCTCT	182		
QY 185	TGATATCGTTCACGAACCTAGAGATATACGTGGCGCTTAATCTCGATTAACGTGGGAGAA	244		
Db 183	TGATATCGTTCACGAACCTAGAGATATACGTGGCGCTTAATCTCGATTAACGTGGGAGAA	242		
QY 245	TTTCGAGAACCTGGAAGAGAAATCCGCGAAGCCAGAGGTACTCTTATGTGACATTTTGGA	304		
Db 243	TGATGTAGAACCTCAAGAAATCCGCGAAGCCAGAGGTACTCTTATGTGACATTTTGGA	302		
QY 305	TGATGTCCCAAGAGAGCTGCCAAATTTATGAACCTAGATCAAGAGTTCCTTTAAAGACA	364		
Db 303	TGATGTCCCAAGAGAGCTGCCAAATTTATGAACCTAGATCAAGAGTTCCTTTAAAGACA	362		
QY 365	CTGTCATACCGATGGAAGAAATACGTTATGTCTTGAAGGAGTGGATCTTTGATGTGAG	424		
Db 363	CTGTCATACCGATGGAAGAAATACGTTATGTCTTGAAGGAGTGGATCTTTGATGTGAG	422		
QY 425	AGACCAAAATGATCATATGATGATTGATATGACACTGMAAGAAAGAGCGCTGATTTTGCC	484		
Db 423	GGACGAAATGACACATGATGATCCGTGTGACAGTGAAGAAAGAGCGATGATTTTGCC	482		
QY 485	TGCAGGATGTACCAACCGCTTTACCTTGGACACCGACACATATATCAAGCAATGCGCAT	544		
Db 483	TGCAGGATGTACCAACCGCTTTACATTTGGACAAAGAAACTATATCAAGCAATGCGCT	542		
QY 545	GTTTGTGGGAGATCTGTTTGGACACCCCTACACACCGTCCCATGACCAT	593		
Db 543	CTTTGTGGGAGACCTGTTTGGACCCCGACACACCGTCCCATGATCAT	591		
RESULT 4				
AW284744/c	546 bp	mrna	linear	EST 19-JUL-2000
LOCUS	LG1_214.H12.g1_A002	Light Grown 1 (LG1)	Sorghum bicolor	CDNA, mRNA
DEFINITION	sequence.			
ACCESSION	AW284744			
VERSION	AW284744.1	GI:6674588		
KEYWORDS	EST.			
SOURCE	Sorghum.			
ORGANISM	Sorghum bicolor			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
AUTHORS	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC			
JOURNAL	clade; Panicoideae; Andropogoneae; Sorghum.			
COMMENT	1 (bases 1 to 546)			
	Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C. and Pratt,L.H.			
	An EST database from Sorghum: light-grown seedlings			
	unpublished (2000)			
	Contact: Cordonnier-Pratt MM			
	Laboratory for Genomics and Bioinformatics			
	The University of Georgia, Department of Plant Biology			
	Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA			
	Tel: 706 542 1860			
	Fax: 706 583 0210			
	Email: mmp@pratt.uga.edu			
	Sequences have been trimmed to exclude PolyA, vector and regions			
	below Pired quality 16. The threshold for highest quality sequence			
	is 20.			
	Seq. primer: 47			
	High quality sequence start: 7			
	High quality sequence stop: 546			
	PolyA-No.			

FEATURES	Source	Location/Qualifiers
Greensstr. 3, D-06466 Gatersleben, Germany		
Email: michael@ipk-gatersleben.de, http://pgrc.ipk-gatersleben.de		
Seq primer: 13 primer for 5' end.		
		1. .700
		/organism="Hordeum vulgare"
		/cultivar="Barke"
		/db_xref="taxon:4513"
		/clone="HM06H13V"
		/clone.lib="Hordeum vulgare Barke roots"
		/flusg.type="roots"
		/lab.host="XLOLR"
		/note="Vector: plasmid pBK-CWV; Site_1: EcoRI; Site_2: XhoI; mRNA was made from roots of spring barley variety 'Barke', a high quality malting variety. Roots were grown for two days on filter paper at room temperature. Cloning sites: EcoRI (5'-end of cDNA) and XhoI (3'-end of cDNA). NOTE: Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter. Average insert size is 1 kb. Sequence trimming: Vector sequences and sequence ends were trimmed from the 5'- and 3'-end until a 50 bp window contains less than two ambiguities. The maximum length was set to 700 bp"
BASE COUNT	204	a 145 c 199 g 150 t 2 others
ORIGIN		
Query Match	48.3%	Score 421.6; DB 9; Length 700;
Best Local Similarity	79.3%	Pred. No. 5,9e-120;
Matches 499;	Conservative 0;	Mismatches 130; Indels 0; Gaps 0;
QY	35	GAAGGTCGCTCGTCTTGCTCCACCCGCAATCCATCGATGAGAAAGAAATTCAGAGTGTAA 94
Db	32	GCAGTTGCGGGGAGACTGCGAGGGGGCGCGGCATGAGACGATCCAGATGGCA 91
QY	95	GACGAGCGGTATGAAAGCATGGTACATGATGATATAGGAAAGACACAGAGCTTCTCA 154
Db	92	GGAGCGAGTATCCAGCATGCTGATGATGATGAGACAGTGAAGAGACCCAGAGGCTTCTCA 151
QY	155	TCACCCGCGAACCAGAAATTCATTCCTGTTGATTAAGCTTACGAACACTGAGATATCAG 214
Db	152	TCACCGTGAAGCCCAAGAGATTCATTCCTTGCAAAACTTTCAGAAATTAAGTGTGTAAG 211
QY	215	CTGAGCGCTAAATCCTGATTAACCTGGAGAAATTTGGAGACCTGAGACATCGCGAAGC 274
Db	212	CTGAGACCTAAATGCTGATTAACCTGGAGAAATTTGGAGACCTGAGACATCGCGAAGC 271
QY	275	CAGAGTTACTCTTATGTGTGACATTTTGATGTGTGCGCCAGAGACAGCTGCCAAATTAGA 334
Db	272	CAGGGAATCTCTTATGTGTGACATTTTGATGTGTGCGGAGAAAGTGGCCAAACTAGCA 331
QY	335	AACATAGATCAAGAGTTCTTGTGAAGAACACCTCATACCGATGAAGAAATACGCTATTG 394
Db	332	GGCCAGCTGAAGAAATTTCTTGTGAAGAACACTTGCATACGATGAAGAGATACGCTATTG 391
QY	395	TCTTGAAGAGGAGGATCTTGTGATGTGAGAGCAAAATGATCAGAGGATTCGATATAGC 454
Db	392	TCTTGAAGGCGATGATCTTGTGATGTGAGAGCAAAATGATCAGAGGATTCGATATAGC 451
QY	455	ACTGAAGAAAGAGGATGATTTGTTCTGCTGACAGGAGATGTACCCGCTTTAAGTTGA 514
Db	452	AGTTAAGAAAGGCGATGATTTGTTCTGCTGACAGGAGATGTATCACCGCTTTAAGTTGA 511
QY	515	CACCGCAACTATATCAAGGCAATGCACTTTTGTGGAGATCTGTTTGAACACCTTA 574
Db	512	TATGAGCAACTATATCAAGGCAATGCACTTTTGTGGAGAGCCCATCTGGAGCGCGTA 571
QY	575	CACCGTCCCGATGACATCTTCGCGCAAGAAAGATTTTGGCTAAACTTCTCAAGTC 634
Db	572	CAATCCCGCCCGATGACATCTTCGCGCAAGAAAGATTTTGTGCAAGAATTTATCAAC 631
QY	635	AGAAGGTGAATAATCAAGCATTTGAAGGCT 663
Db	632	AGANTGGGAGCCCAAAACCGTTGCGAGGCT 660

[illegible]

Oy	488	AGATGATCCACCGCTTTAGCTTGGAACCGCAACAATATATCAAGCAATGGACTGT	547
Db	624	AGGAATGTATCATCAGCTTACCTTGATATGTGACACTACTACATAGSAAATGGCTTANT	683
Oy	548	TGTTGGCATCTCTGTTTGGACACCCTACACCGTCGCCATATGACCATCTCTCTGCAGAANA	607
Db	684	TGTGGGTACCGCTGTGTGGACACCTTAAATGCTTCCCATATGACCATCTCTCCGGATTGGA	743
Oy	608	GGAATTTTTGGCTAATCTCTCAATCTCAAGAGTGGAATAATCAAGCAATGTGAAGGCTTCTG	667
Db	744	GGAGTATCTGAGAAATATCATCAA---CCGAGGTGGAACCAACGAGCGTAGAGCTGTG	800
Oy	668	AGGTTT 674	
Db	801	AGCTGT 807	
RESULT 7			
LOCUS	C28663	467 bp	mRNA linear EST 04-APR-2002
DEFINITION	C28663 Rice callus cDNA Oryza sativa (japonica cultivar-group) cDNA		
ACCESSION	C28663		
VERSION	C28663.1	GI:2312508	
KEYWORDS	EST.		
SOURCE	Oryza sativa (japonica cultivar-group).		
ORGANISM	Oryza sativa (japonica cultivar-group). Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Eriactoidae; Oryzaceae; Oryza. 1 (bases 1 to 467) Yamamoto,K. and Sasaki,T. Rice cDNA from callus 1997 Unpublished (1997) Contact: Takuji Sasaki National Institute of Agrobiological Resources Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan Tel.: 81-298-38-7441 Fax: 81-298-38-7468 Email: tsasak@abrr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/ PROJECT = RGP.		
FEATURES			
source	Location/Qualifiers		
	1..467		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/cultivar="Nipponbare"		
	/db_xref="taxon:39947"		
	/clone="C61979_1A"		
	/clone_1lb="Rice callus cDNA"		
	/tissue_type="callus"		
	/dev_stage="calulus"		
	/note="Organ: panicle; Rice cDNA from panicle at ripening stage"		
BASE COUNT	150 a 92 c 119 g 104 t	2 others	
ORIGIN			
Query Match	48.1%; Score 419.6; DB 14; Length 467;		
Best local Similarity	97.2%; Pred. No. 2.le-119;		
Matches 447; Conservative	0; Mismatches 11; Indels 2; Gaps 2;		
Oy	1	AGAGACAAAATAAACGAATCCATGCGCATATCGAAGTTCGCTTGTCTTCACCCCG	60
Db	8	AGAGCAACAAAAACGAATCCATGCGCATATCGAAGTTCGCTTGTCTTCACCCCG	67
Oy	61	CAATCCCATGAGAGAACGAATTCACGAGATGTAACAGGAGTATGTAAGCAATGGTATCA	120
Db	68	CAATCCCATGAGAGAACGAATTCACGAGATGTAACAGGAGTATGTAAGCAATGGTATCA	127
Oy	121	TGG-ATGATAGGAGAGGACGAGAGGCTTCTCTATCACCGCGAACCAGCAATTCATT	179
Db	128	TGGAAATATATGCGAAGAGGACGAGAGGCTTCTCTATCACCGCGAACCAGCAATTCATT	187

[illegible]

Accession	Version	KeyWords	Source	Organism	Title	Journal	Comment
AM678360	1	GI:7552089	Sorghum bicolor	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.	Unpublished (2000)	Contact: Cordonnier-Pratt MM	Laboratory for Genomics and Bioinformatics The University of Georgia, Department of Plant Biology Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel: 706 542 1860 Fax: 706 583 0210 Email: mmpratt@uga.edu
REFERENCE	AUTHORS	Cordonnier-Pratt,M., M., Gingle,A., Marsala,C., Sudman,M. and Pratt ,L.H.					
		An EST database from Sorghum: water-stressed plants					
		Seq primer: JEN REV					
		High quality sequence stop: 512					
		PolyA-No.					
FEATURES	source	location/Qualifiers					
		1..574					
		/organism="Sorghum bicolor"					
		/db_xref="taxon:4558					
		/clone_lib="Water-stressed 1 (Wst)"					
		/note="Organ: Mix of 5-week old plants on days 7 & 8 after water was withheld; Vector: lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were prepared by mass excision."					
BASE COUNT	166 a	126 c	144 g	138 t			
ORIGIN							
Query Match	47.9%	Score 418:	DB 10:	Length 574:			
Best Local Similarity	87.5%:	Pred. No. 7.2e-119;					
Matches 457:	Conservative	0;	Mismatches 65;	Indels 0;	Gaps 0;		
OY	135	GAGACCCAGAGGCTTCCTCATCACCGCGCAACCAGAAAGATTCATTCTGTGGTAAAGCTT	194				
Db	9	GAAACACCGAGGCTTCCTCATCACCGTGATCCCAAAGAATTCATTCTGTGAATGACTTT	68				
OY	195	ACAGAACCTAGAGTAATACGACGCGGCCCTAAATCCTGATACTGGAGATTCGCAGAAC	254				
Db	69	TCAAACTGSGAATCATCACTGCGCCCTGAATCTCGACAATCGGAGATGATGAGAAC	128				
OY	255	CTGAGAGAAATCCGCGAGGCAAGAGGTACTCTTATGTGGACATTTGTGATGTGTGCCCA	314				
Db	129	CTCAAGAAAATCCGTGAAGCCAGAGGTACTCTTACATGACATTTGTGATGATGATGCCA	188				
OY	315	GAGAAGCTGCCAATATATGAACCTAAGATCAAGACTTTCTTTGAAGAAGACCTGCATATAC	374				
Db	189	GAGAAAGCTGCCAATATATGAGATTAAGATAAAGATTTCTTTGAAGAAGACCTGCATATAC	248				
OY	375	GATGAAGAAATPACGCTATTGTCTGTGAAGGAGTGGATCTTGGATGTGAGAGACCAAAAT	434				
Db	249	GATGAAGAAATPACGCTATTGTCTGTGAAGGAGTGGATCTTGGATGTGAGAGACCAAAAT	308				
OY	435	GATGATGATGCTGTATGACATCGAAGAAAGAGGACATGATGTTGTCTGCCGTCGAGGATC	494				
Db	309	GACCAAGTGATTCGCTGTAGACATGGAAGAAAGAGGACATGATGTTGTCTGCCAGGATG	368				
OY	495	TACCAACGCTTTACGTTGGACACGCAACATATATCAAGGCAATGCGCATGTTGTGGGC	554				
Db	369	TATCACCGCTTACATTTGGACATGAGAACATATATCAAGGCAATGCGCTTTGTGGGC	428				
	555	GATCCTGTTTGGACACCTCTCAACGCTCCCATGACATCTTCTCGCAAGCAAGAGCTTT	614				

Db	429	GACCCTGTTGGACGCCGCTACACCGTCGCCATGATCATCTGCACGACGAGACGAGTAT	488
Oy	615	TTGGCTAAACTTCTCAGTCACGACGAGTGAATTAANTCAGCAGT	656
Db	489	CTGGACAGACGCTCTCAAGCCAGACGACGACCTTTGACGT	530
RESULT 10			
BI542476		600 bp	linear
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

```

Db      136 CAAACTTCCGACTAGCAATATTAACTGGGCGCTAAATGCTGATGATGGGGAATGA 195
QY      248 CGAGAACTGTAAGAAATCCGCAAGCCAGAGGTTACTCTTATGTGACATTTGTGATGT 307
Db      196 TTAGAACTCAAGAAATTCGTGAGCCAGGGGTTACTCTTATGACATTTGTGATGT 255
QY      308 GTGCCAGAGAGCTGCCAAATATGAACTTAAGATCAAGACTTCTTTGAAGAACCT 367
Db      256 GTGCCAGAGAAATGGCAAAATTAAGAGGCTAAAGATTAAGATTTCTTTGAAGAACCT 315
QY      368 GATATCCGATGAAGAAATAGCGTATTTCTTGAAGAGAGTGGATCTTTATGTGAGAGA 427
Db      316 GATATCTATGAAGATAGATGATATGCTTGAAGAGAGGCTGGATCTTTATGTGAGAGA 375
QY      428 CCAAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 487
Db      376 CCAAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 435
QY      488 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 547
Db      436 AGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 495
QY      548 TGTGGGATCTCTTTGGACACCTTGAAGAGGCTGGATGATGATGATGATGATGATGATG 607
Db      496 TGTGGGATCTCTTTGGACACCTTGAAGAGGCTGGATGATGATGATGATGATGATGATG 555
QY      608 GGAGTTTGGCTAACTCTCAAGTCAGAGG 640
Db      556 GGAGTACGTGAGAAATCATCAACGAGGTGG 588

RESULT 11
LOCUS   BG343468 978 bp mRNA linear EST 22-OCT-2001
DEFINITION
HVS690005N13f Hordeum vulgare pre-anthesis spike EST library
HVCDA0008 (white to yellow anther) Hordeum vulgare cDNA clone
HVSME90005N13f, mRNA sequence.
ACCESSION BG343468
VERSION BG343468
KEYWORDS
SOURCE EST.
ORGANISM Hordeum vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 978)
AUTHORS Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Fritsch, D., Yu
, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton
, R.D., Close, S.J., Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex pre-anthesis spike cDNA library
Unpublished (2001)
JOURNAL Contact: Ming RA
COMMENT Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: twing@clemson.edu
Total hg bases = 505
Seq primer: AATTAACTCTCACTAAAGG
High quality sequence stop: 879.
FEATURES
source
1..978
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSME90005N13f"
/clone_lib="Hordeum vulgare pre-anthesis spike EST library
HVCDA0008 (white to yellow anther)"
/tissue_type="pre-anthesis spike"
/lab_host="SOLR"

```

```

/note="Vector: lambdaZAP, Site_1: EcoRI, Site_2: XhoI;
Plants were grown in the greenhouse at the University of
California, Riverside (Fenton, SJ Close, TJ Close). Whole
spike with awns trimmed were collected at white, green and
yellow anther stages (Fenton). Total RNA was prepared from
each pool, equal quantities of all three RNA pools were
combined, poly(A) RNA was purified from the mixture, one
primary unamplified cDNA library was made, and 1 million
pfu were in vivo excised to give plasmidic SW(-) cDNA
phagemids. These steps were performed in the TJ Close lab
(Chol) at the University of California, Riverside.
Phagemids were plated and picked at the Clemson University
Genomics Institute (CUGI) (Begum, Palmer, Fritsch, Atkins
and Wing) Plasmid DNA preparations, DNA sequencing and
sequence analysis were performed at CUGI (Wing, Yu, Fritsch
, Henry, Simmons, Oates, Rambo, Main). The sequence has
been trimmed to remove vector sequence and contains a
minimum of 100 bases of phred value 20 or above. For more
details on library preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and Physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/g9pages/bgn/31/cover.html)"
BASE COUNT 275 a 207 c 275 g 219 t
ORIGIN
Query Match 47.9%; Score 417.8; DB 12; Length 978;
Best local similarity 80.5%; Pred. No.1e-118;
Matches 488; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 35 GAAGTTCCTCTCTTCCACCCCGCATTCACATGAGAGAGCAATTCAGATGTAA 94
Db 18 GCGATTGCGCGGAGACGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 77
QY 95 GACGGAGGATAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 154
Db 78 GGAGCAGGTCATCCAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 137
QY 155 TCACCGCGAAGCAAGATTCCTGTTGATGATGATGATGATGATGATGATGATGATG 214
Db 138 TCACCGCGAAGCAAGATTCCTGTTGATGATGATGATGATGATGATGATGATGATG 197
QY 215 CTGGCCCTTAATCTGTAACTGGAGAAATTCGAGAACCTGAGAGATCCGGAAC 274
Db 198 CTGGAACTTAATCTGTAACTGGAGAAATTCGAGAACCTGAGAGATCCGGAAC 257
QY 275 CAGAGGTTACTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 334
Db 258 CAGGGAATCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 317
QY 335 AACTAGATCAAGATTTCTTTGAAGAACCTTGCATCTGATGAAGATAGCGTATG 394
Db 318 GGCACAGCTGAAGATTTCTTTGAAGAACCTTGCATCTGATGAAGATAGCGTATG 377
QY 395 TCTTGAAGGAGTGAATCTTGAAGAACCTTGAAGAACCTTGAAGAACCTTGAAG 454
Db 378 TCTTGAAGGAGTGAATCTTGAAGAACCTTGAAGAACCTTGAAGAACCTTGAAG 437
QY 455 ACTGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 514
Db 438 AGTTAAGAAAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG 497
QY 515 CACCGCAACTTATCAAGGCAATGCGATGTTGTTGGCATCTGTTGGACACCTTA 574
Db 498 TACTGAAGCACTATCAAGGCAATGCGATGTTGTTGGCATCTGTTGGACACCTTA 557
QY 575 CAACGCTCCCGATGACATCTTCTCGAAGAAAGAGATTTGGCTTAACCTTCAAGTC 634
Db 558 CAATCCCGCCATGACATCTTCCAGCTAGAAAGAGATATGTCACAGATATCAACAG 617
QY 635 AGAAGG 640

```


TITLE Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
JOURNAL Development of Barley Transcriptome Resources
COMMENT Unpublished (2001)
On Jan 10, 2002 this sequence version replaced g1:18113748.
Contact: Maugh R, Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: estescr@sari.ac.uk
All sequence has a phred quality score of 20 or over
Seq primer: M13 reverse.

FEATURES

source

Location/Qualifiers
1. .709

/organism="Hordeum vulgare"

/cultivar="Optic"

/db_xref="taxon:4513"

/clone="EBro08_S0003_P14"

/clone_lib="root", 3 week, drought-stressed, cv Optic,
EBro08"

/tissue_type="root"

/dev_stage="3 week"

/lab_host="DH10B"

/note="Vector: pSPORT1, Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from roots of 3 week old drought stressed barley
plants. Developed as part of the barley transcriptome
resources of BBSRC/SEERAD funded cereal ICF (Investigating
Gene Function) project."

BASE COUNT

206 a 149 c 196 g 158 t

ORIGIN

Query Match

47.8%: Score 417.2; DB 13; Length 709;

Best Local Similarity 80.5%; Pred. No. 1,4e-118;

Matches 488; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 35 GAAGGTCGCTCTTCTTCACCCGCAATCCAGATGAGAAATTCAGATGCTAA 94
DB 18 GCAGTTCGCGGAGACTGCGAGGGCCGCCGATGAGAAAGATTCAGATGCA 77
QY 95 GACGAGGTGATAGAGATGATGATGATGATGATGATGATGATGATGATGAT 154
DB 78 GGACGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 137
QY 155 TCACCGGCAACCAAGAAATTCCTGCTGATAGCTTACAGACTAGAGATGATGAT 214
DB 138 TCACCGGCAACCAAGAAATTCCTGCTGATAGCTTACAGACTAGAGATGATGAT 197
QY 215 CTGGCGCTTAATCCTGATACTGGGAGAAATTCGAGAACTGAGAAATTCGCGAAGC 274
DB 198 CTGGAACCTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 257
QY 275 CAGAGGTACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 334
DB 258 CAGGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 317
QY 335 AACTAGATCAAGATTCCTTGAAGAACACCTGATACCATGAGAAATACCTATG 394
DB 318 GGCAGAGCTGAAGATTCCTTGAAGAACACCTGATACCATGAGAAATACCTATG 377
QY 395 TCTTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 454
DB 378 TCTTGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 437
QY 455 ACTGAAGAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 514
DB 438 AGTTAAGAAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 497
QY 515 CACGCAACTATATCAAGCAATGCGACTGTTGTTGGCGATCCTGTTGGACACCTTA 574
DB 498 TAGTGACAATATCAAGCAATGCGACTGTTGTTGGCGATCCTGTTGGACACCTTA 557

QY 575 CAACGCTCCCATGACCATCTCTGCAAGAAAGAGTTTGGCTAAACTCTCAAGTC 634
DB 558 CAATGCCCCCATGACCATCTCTGCAAGAAAGAGTATGTGACACATTAATCAACAG 617
QY 635 AGAAGG 640
DB 618 AGGTGG 623

Search completed: December 1, 2002, 21:40:09
Job time : 2194 secs

7

.

ge

